DOCUMENT RESUME

ED 184 808 SE 029 669

AUTHOR Piel, E. J.; Truxal, John G.
TITLE Man and His Technology: Problems and Issues.
INSTITUTION State Univ. of New York, Stony Brook. Coll. of Engineering.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Jan 73
NOTE 269p.; Not available in hard copy due to marginal legibility of original document.

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Data Analysis; *Decision Making; Energy; Environment; Environmental Education; Evaluation Methods; Natural Resources; Pollution; *Population Growth; *Public Health; Public Policy; Sciences; *Technological Advancement; *Technology

ABSTRACT

This book looks at a sampling of notable socio-technical problems within a rational context. It seeks to determine if problems exist with recent decisions in the application of technology; how threatening any problems may be; the source of problems; and examines alternatives available to the nation. The broad goal of this rational consideration is to seek alternative applications of technology, the alternatives having been identified by rational technology assessment. (Author/RE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Mary L. Charles
of the NSF
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

MAN AND HIS TECHNOLOGY:
PROBLEMS AND ISSUES

E. J. Piel

John G. Truxal

College of Engineering

State University of New York at Stony Brook

ED 029 669

PREFACE

An Arabian sultan became thoroughly disgusted with his two sons; neither seemed fit to inherit his power and wealth, as they were interested only in girls and fast horses. When he died, his will ordered that each son should select his fastest horse and race to Mecca. The son whose horse reached Mecca last would inherit everything.

The two young men selected their horses and started out. After a week they were still only a mile from home, as each dawdled in every possible way. It was clear to both that they would spend the rest of their lives on the trip, and no one would enjoy the riches.

In desperation, they consulted a wise man, who listened to their story and thought briefly. He then said to them that he would solve their problem if they did exactly as he ordered. When they agreed, he spoke briefly to them, whereupon they each dashed from the room, jumped on the horses, and galloped full speed toward Mecca.

What had he told them? The wise man had studied the problem and realized that the trouble arose because the winner was the horse to take the maximum time to Mecca. To convert this to a minimum problem, he ordered each young man to ride his brother's horse. Now the goal of each rider would be to reach Mecca as soon as possible.

This story, which really has nothing whatsoever to do with technology, illustrates one of the basic objectives of this book. In the following chapters, we will consider a sequence of socio-technological problems (health services, auto accidents, pollution, energy shortage, and so on). In each of these problem areas, the United States (and indeed much of the world) seems to be in a nearly hopeless situation. It is not clear how governmental or social action can improve the situation.

In the energy crisis, for example, building more electricity energy generating stations will increase air and other pollution, further deplete already scarce natural resources, and introduce safety problems. The only two alternatives seem to be more costly sources or decreasing per capita use. In the former approach, any rise in the cost of electricity in our cities drives more manufacturing industry to rural areas (and the welfare problem worsens) or weakens the international trade position of this country. If we try to decrease usage, the only effective means seem to involve governmental restrictions on personal freedom.

As the country selects among the alternatives which are open to us, intelligent decisions depend upon public technological literacy. The decision makers and the public must be able to look at the problem, understand the factual core of knowledge about this area, and appreciate the significance of the various alternatives. While many people agreed with the 1971 decision in Congress on the SST (the supersonic passenger airplane), few could agree with the hysteria and misinformation which beclouded what should have been an informed public debate.

Thus, the first objective of this book is to attempt to look at a sampling of important, socio-technological problems in a calm and rational way: to consider first whether there actually is a problem and, if so, how bad it is; next to determine why the problem has arisen; and finally, to see what alternatives are available to us as a nation.

The broad goal of such rational consideration is to seek an alternative (a plan of action) which benefits all interested parties. In so many of the environmental and social issues of the past few years, the problem has been viewed simplistically as just a conflict between industry and human values, between the government and the people, or between technology and the environment. In many cases, such a head-on collision emerges merely because the problem is simplified so much that essential elements are omitted entirely from consideration.

Again, a classic story illustrates the ridiculous position that can arise when seemingly competitive forces refuse to cooperate. Two ice cream vendors wheel their carts to a mile-long beach each morning. There is no reason for the public to prefer one over the other, so the customer goes to whichever one is closer. Where should the one who arrives first set up his stand? How about the one who arrives later?

The first one should locate in the middle of the beach, the second immediately beside him. Thus, each day the two are shoulder-to-shoulder in the center of the beach. If either one locates anywhere else, the other will have more than half of the customers.

Clearly, this is a ridiculous way to act. They each would do better (and the public would be better served), if they cooperated and each located 1/4 mile from an end of the beach. Then the two vendors would be a half-mile apart, but no potential customer would be more than a quarter mile from a vendor. And each would service half the beach.

This solution, so logical to an outsider, requires that the two vendors cooperate. They must discuss their problem of where to locate, reach a rational and thoughtful decision, and finally both abide by that decision. In the same way, the conflicting forces in the much more complex, socio-technological problems must, for the public good, compromise on the best possible policy even if it does not correspond to the maximum potential profit for the company or the absolutely minimum damage to a part of the environment. Traffic fatalities could be stopped abruptly by forbidding any vehicles, but the social and economic upheaval would be catastrophic; on the other hand, the failure to require safety measures which could save thousands of lives represents an equally indefensible approach.

As problems become more complex and as they impact more strongly on the quality of human life, the public and the decision-makers must approach these problems with a greater depth of understanding. This is the background motivation for this book, which is based upon courses given by the authors to general undergraduates and to adult, continuing education, graduate students, both at the State University of New York at Stony Brook. This book has evolved from the authors' work over the past seven years on the Engineering Concepts Curriculum Project, which developed the secondary school text The Man-Made World. The authors express their gratitude to their colleagues on the ECCP activity, and recognize the support provided to ECCP by the National Science Foundation. The development of the college course has also been aided by grants from the U. S. Steel Foundation, IBM, Public Service Electric and Gas Corp. of New Jersey, the TRW Foundation, and the Esso Education Foundation.

E. J. Piel
J. G. Truxal

Stony Brook, New York
January 10, 1973

TABLE OF CONTENTS

Chapter 1 SYSTEMS AND DECISIONS 1

1. System Studies 4
2. Precipitation Control 8
3. Natural Disasters 13
4. Conclusion 14

Chapter 2 POPULATION 16

1. Inertia in the Population Problem 16
2. Sea Lamprey in the Great Lakes 19
3. Buffalo 22
4. U.S. Population 28
5. Use of the Population Model 33
6. Exponential Growth 35
7. Graphs of Exponential Growth 40
8. Possibility of Population Control 45
9. Final Comments 51

Chapter 3 DELIVERY OF HEALTH SERVICES 53

1. Health Services 54
2. The Health Services System 56
3. Current Performance of the U.S. System 58
4. Reasons for Poor U.S. Performance 64
5. Male Health 70
6. Central-City Health Problems 74
7. Nutrition 79
8. Management of Health Delivery 85
9. Conclusion 87

APPENDIX 3-1 FEDERAL IMPLEMENTATION 89

Chapter 4 EMERGENCIES 96

1. Automobile Accidents 97
2. Auto-Safety Controversy 105
3. Alcohol and Driving 117
4. Emergency Medical Service 126
5. Other Accidents 132
6. Conclusion 135

Chapter 5 DECISIONS 139

1. Drug Control as a Decision Problem 140
2. Complexity of the Drug-Control Problem 145
3. Importance of Information 150
4. Simple Choices 157
5. One Man, One Vote 164
6. Competitive Decision-Making 170
7. Decision Problems Without Solutions 179

Chapter 6 THE ENERGY CRISIS 183

1. What is Energy 184
2. Problems of the Energy Crisis 187
3. U.S. Energy Use 189
4. Contributory Factors 192
5. Alternatives or Options 196
6. New Technology 201
7. Final Comment 207

Chapter 7 NOISE ENVIRONMENT 209

1. Measuring Noise 210
2. Frequency 211

3. Amplitude 214
4. Other Characteristics of Noise 217
5. Physiological Effects of Noise 221
6. Psychological Effects 224
7. Transportation Noise 226
8. Other Sources of Urban Noise 229
9. Conclusion 231
Chapter 8 EPILOGUE 233
Problems
Index

CHAPTER 1

SYSTEMS AND DECISIONS

September, 1976. A major hurricane at sea is heading directly toward the densely populated New Jersey-New York region. Within the storm, wind velocities of 100 miles per hour have been measured by observation planes. Radio and television are broadcasting frequent warnings to area residents, but there is very little they can do. The storm promises to be fully as severe as Betsy in 1965 and Camille in 1969, each of which caused injuries and deaths in addition to property damage of $1.5 billion. The winds were the primary cause of destruction, although there was also some flooding.

Meanwhile, in a Washington office a man debates whether to try to modify the storm by cloud seeding. Most meteorologists are convinced that the maximum wind velocity of a hurricane can be reduced at least slightly by dropping massive amounts of silver iodide crystals in the outer eyewall region. The obvious decision is to seed.

For several reasons the decision is not so simple. First, there is really very little scientific knowledge about the probable effects of seeding. Only two hurricanes have been seeded in the past, with only one of the experiments really useful as a source of knowledge. On August 18 and 20, 1969, hurricane Debbie was seeded, with reductions of 31% and 15% in the peak winds after the crystals were dropped. In 1971, Ginger was seeded, but in this case the silver iodide was placed in the heavy-rain region well outside the eyewall; the seeding had no significant effect.

Since 1971, there have been no experimental seedings even though the program, called Project Stormfury, has continued. Why no experiments for five years? The Project has been working under the rules that no seeding is allowed if the storm threatens a coastal area or may hit a populated area within 18 hours. So little is known scientifically about the effects of seeding on hurricanes, Project Stormfury has been avoiding modification of any storm which might immediately affect people in large numbers. Hence, experiments have been restricted to a relatively small area of the Atlantic Ocean south of Puerto Rico -- a region close enough to scientific instruments to allow detailed measurements of storm characteristics, but far enough from populated areas so that, if the seeding should cause intensification of the storm, there would be no direct effects.[1]

The decision of "our man in Washington" is further complicated. The maximum wind velocity is now 100 miles/hour, with the center of the storm 100 miles off shore (or 12 hours before it hits the New Jersey-New York coast). Past experience with hurricanes indicates that this peak wind will probably increase as the storm moves across the open water, although there might be a decrease. By the time it hits land, the maximum wind velocity is likely to be 125 miles/hour. If the hurricane is seeded, the probable or hoped-for effect will be a maximum wind velocity of 110 miles/hour at landfall (Fig. 1-1). This reduction of peak wind from 125 to 110 (if seeding is completely successful) will result in a 40% decrease in the damage.[2] In dollar terms, $400 million will be saved if the damage without seeding will be one billion dollars; this is a strictly economic accounting which does not take into consideration the human misfortune.

---

[1] The instrumentation used in hurricane studies is described in the article: James W. Meyer, Toward Hurricane Surveillance and Control, Technology Review, October/November 1971, pp. 59-66.

[2] In mathematical terms, the expected damage is proportional to the maximum wind velocity raised to the 4.4 power.

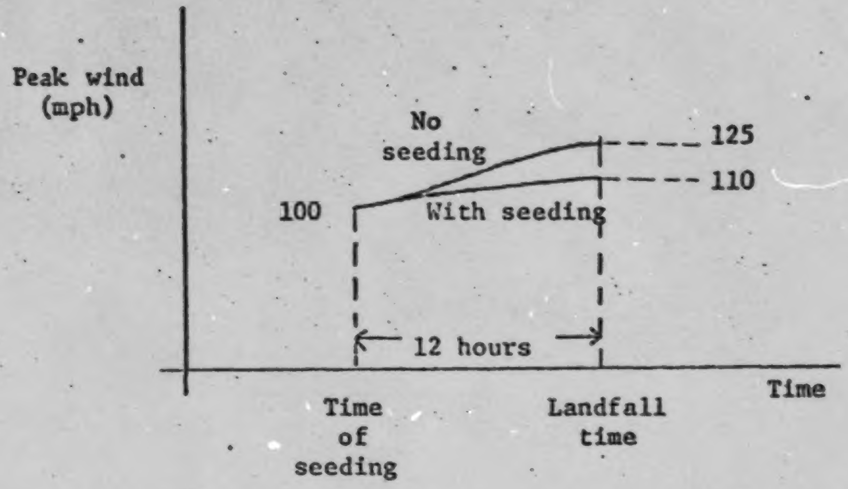


Fig. 1-1 Possible change in peak wind as hurricane approaches land

The problem confronting the "man in Washington" is now clear. If he orders seeding and it is successful, the public will learn that the peak winds increased from 100 to 110 mph after the seeding operation. Since there is a small chance the peak wind might have decreased before landfall, the natural public conclusion will be that the seeding intensified the winds and increased the damage. If he does nothing (no seeding), the storm is likely to hit land with peak winds of 125 mph, with catastrophic effects. Yet under this latter option, there will be no strong criticism of the government, since the public can be told with considerable honesty that too little is known scientifically about hurricane seeding to justify this action. Finally, the decision is further complicated by the fact this is an election year; the administration, which will be held responsible by the public for seeding results, must face the electorate in less than two months.

Comment

This story, which is likely to happen in the next few years, is a typical *socio-technological problem*.[1] The *technological* refers to the strong content of applied science and engineering: the problem arises only because of the progress which has been made during the last two decades in weather observation and modification. The *social* element enters because the technology has such a direct and strong impact on a segment of society.

The problem is also typical because it requires a *decision* which is strongly influenced by technological, economic, political, and social factors. The *optimum* decision has to be determined from not only scientific and economic considerations, but also public attitudes.

The problem is typical, too, because a decision *must* be made. There is no such alternative as postponement until more scientific knowledge is available or until a public education effort can be completed. The decision presents only two options: seed or do not seed. No decision is a decision not to seed. In the same way, although often less dramatically, whether to allow a new drug to be sold, a new nuclear energy plant to be built, or a new subway system to be started is a decision which is forced on the government and the people. Postponement merely means that the decision has been negative for the period of delay.

1. System Studies

Through the following chapters we consider a small selection of socio-technological issues similar to the hurricane-seeding problem. The prime goal is an understanding of the basic approach to such problems which comes under the title of system studies.

---

[1] A lengthy mathematical analysis of the hurricane-seeding decision problem is summarized in the article by R.A. Howard, J.E. Matheson, and D.W. North, The Decision to Seed Hurricanes, *Science*, June 16, 1972, pp. 1191-1202.

In a system study of a particular area, we attempt first to determine whether a problem really exists. If so, what are the origins of the problem? Why has it risen to prominence in recent years? What current factors are aggravating the problem?

The more quantitative we can be in answering these questions, the better understanding we can achieve of the nature and magnitude of the problem. In the language of today's physical or social scientist, we seek a model of the system or problem: a statement in quantitative terms of those aspects of the situation which are relevant.

This model is almost always dynamic or changing with time. It is this dynamic character which makes the problem interesting and challenging. The dynamic model permits a prediction of the future situation, so that action can be taken now to modify that future toward more attractive directions.

The importance of the dynamic model is illustrated by all of the case studies throughout the remaining chapters. Here an almost trivial example serves to illustrate this fundamental concept.

In 1956 a few African bees were transported to Brazil to breed with the European bees already there. The hope was to develop a more productive strain. A year later, 26 bee swarms, headed by African queens, were accidentally freed and the result of native cross-breeding began slowly to spread throughout South America. Unfortunately the result of the racial mixture is a variety which unexpectedly becomes exceedingly aggressive, attacking animals and people (horses seem to be particularly susceptible).

The danger is that, if this variety reaches the United States, widespread public reaction will result in the sharp curtailment of the bee-keeping industry. The consequences would be a serious blow to agriculture as well as horticulture, where pollination by bees is so important.

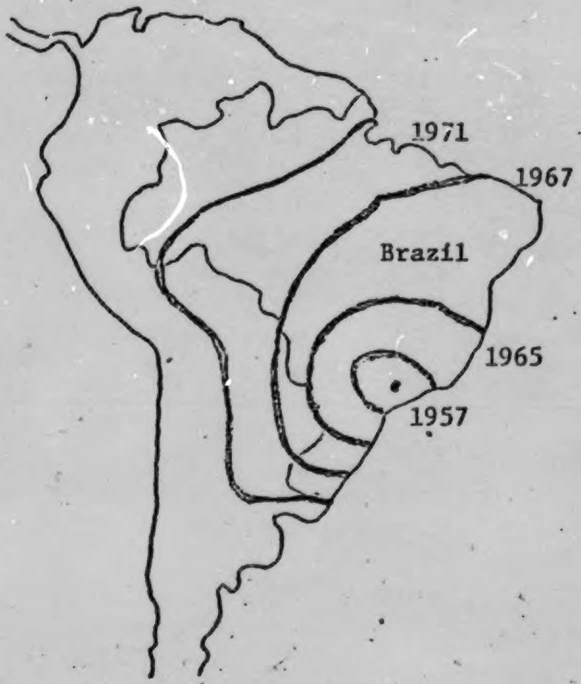


Fig. 1-2 Spread of the Brazilian bee

Understanding of the problem depends on a dynamic model which shows the way the new Brazilian bee is spreading (Fig. 1-2). This model shows the region in which efforts should be focussed to stop the spread, with the particular area dependent on the time scale involved in the creation of barriers. For example, the model shows where immediate quarantine measures might be useful. If scientific studies reveal that it is possible to modify genetically the strain, the model indicates where such efforts should be centered. Another alternative is to saturate a barrier area with a passive race of bees, so that queens in the area would have only a low probability of mating with the aggressive, Brazilian males.[1]

Options and alternatives

Once the quantitative model is obtained for a particular problem, the second step is to determine the options or alternatives which are available for amelioration of the situation. In the hurricane-seeding problem, there are only two options, to seed or not to seed. In the Brazilian-bee problem, we listed several possibilities for positive programs to arrest the spread.

More generally, there may be a very large number of alternatives, each having both advantages and disadvantages. Decision-making involves determination of that single strategy or the mix of strategies which seems likely to have the maximum positive value. Here again, difficulties arise because any strategy is likely to have not only quantifiable effects (for example, economic costs and benefits), but also human and social effects with which it is very difficult to associate numbers. If a proposed nuclear energy plant can be located at three different sites, it is often easy to compare the three costs of electricity which will result from different land, construction, operation, and transmission costs. It is much more difficult to determine the social and human values of lower-cost electricity in terms of such factors as added industrial growth and more jobs. It is

---

1. A detailed study of this problem is available in the "Final Report of the Committee on the African Honey Bee," June 1972 available from the Administrator, Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C. 20250.

equally difficult to estimate the relative costs of thermal pollution and radiation or safety hazards.

Again, the decision must be made, for no decision is really a decision not to build the plant and, hence, to limit the public use of electrical energy.

In this system approach, man and society are viewed as parts of both nature and the man-made processes and world with which man interacts dynamically. The human control of the development of technology requires that the models used for decision-making must include the human and social factors on which that technology impacts. And the models must be sufficiently complete to permit the best possible estimates of the future effects, however indirect, of that technology.

2. Precipitation Control

Project Stormfury, designed to study hurricane seeding, is just a small part of the total weather modification effort in the United States. The program is under a number of different federal agencies and departments, as well as state governments. There are two principal objectives: to improve the level of precipitation, rain or snow, in selected regions; and to decrease the effects of natural disasters. This section summarizes the first, and the second phase is discussed in Section 3.

The control of rainfall and snowfall is primarily directed toward increasing precipitation in dry regions or where water is scarce, although there has been some experimentation designed to reduce snowfall in areas where transportation is difficult or snow removal costs heavy.

History

Average annual rainfall varies from zero over some of the driest deserts to 460 inches at one site in Hawaii, with the usual between 20 and 60 inches. From the earliest times, man has tried to increase rainfall in times of drought. By the first part of the 1800's man recognized that the clouds should be perturbed, and cannonballs were shot into the air to shake the moisture loose.

The first federal program was in 1891, when dynamite was lofted to the clouds in balloons. These unsuccessful attempts were followed by experiments mixing various chemicals with smoke.

On November 13, 1946 the General Electric Company scientists, Irving Langmuir and Vincent J. Schaefer, seeded clouds over Mt. Greylock, Massachusetts with dry ice crystals. Minutes later, snow fell in the first successful experiment. Subsequently, Bernard Vonnegut (now of the State University of New York) discovered that silver iodide crystals were effective, and the foundation was laid for the hundreds of projects now carried out annually in the United States.

The effectiveness of cloud seeding to increase precipitation has only been demonstrated in the past few years, as the scientific explanation of rain has evolved. In the early days, seeding had uncertain results, with rainfall perhaps decreasing as often as it increased. Recent studies have led to complicated computer models derived from extensive measurements of the physical characteristics of clouds: the temperature distribution, the density of water vapor, wind velocities, pressure, and so forth. From these models, predictions on the effect of seeding can now be made with considerable confidence.

For example, meteorological studies showed that the temperature in the upper part of the cumulus cloud is critical. If this temperature is above -20°F, silver iodide crystals serve as particles around which water vapor can condense to form ice crystals which grow and fall. If the temperature is below -24°F, the silver iodide tends to prevent ice formation, and precipitation is reduced. This is a vastly oversimplified explanation, but it illustrates the type of scientific knowledge which has led to today's successful rainmaking.

Current activities

Two programs have received widespread publicity. The first is Project Skywater, in which seeding of the clouds above the San Juan Mountains of Colorado is proposed to increase the snowfall. The potential result is a 16% increase in the spring runoff water feeding the Colorado River, which ultimately provides water to Nevada, Arizona, and southern California (some 60% of the water for the Los Angeles area, for instance). Small-scale experiments have indicated that an expenditure of $5 million for instrumentation, seeding, and monitoring could yield additional water valued at more than $100 million.

The second large-scale project recently was in 1971 in Florida, when the state and the National Oceanic and Atmospheric Administration (NOAA) joined forces to seed promising rain clouds every other day for two months. The drought was broken, with 180,629 acre feet of rain drenching an area of 3000 square miles; nearly 2/3 of the rainfall was credited to the seeding. Just in terms of the value of the water, the $165,000 experiment gave a benefit of $5 million. In addition, fires in the Everglades were put out, with an unmeasured saving in natural resources.

Other activities include a large state-operated seeding program in central Oklahoma and experiments in the eastern Great Lakes region to reduce the snowfall.

Decision questions

The recent success of precipitation-control experiments raises a host of questions in public policy which will increasingly appear in public discussions in the next few years. There is by no means general popular support for such weather modification. While the increased rainfall may be generally welcome in the region where the seeding is done, there are often strong objections:

(1) Public opinion studies have been made which show strong opposition, often from groups with deep religious convictions or individuals poorly educated. People in the northeast seem particularly opposed.

(2) Environmentalists are seriously concerned that increased precipitation may alter the balance among plants and animal species. Certain weeds, pests, and diseases may flourish. In a Bureau of Reclamation project in Montana, complaints arose that the increased snowfall interfered with normal migration of the elk.

(3) At least a minority of the population in a region are always adversely affected. For example, in the Colorado snow augmentation program, how can the increased value of the Colorado River water be used, at least in part, to compensate the Colorado residents for the difficulties and expenses associated with 25% more snow each winter?

(4) With the prevailing weather motion west to east in the United States, the question arises: what happens to precipitation eastward of the seeding area? The argument is usually made that less than 10% of the

water in the air falls naturally; consequently, a slight increase in one area will not seriously affect precipitation to the east. The very incomplete understanding of weather dynamics, however, really leaves the question unanswered.

As an extreme example of this problem, rainfall in the middle Atlantic states in 1972 far exceeded the record for over a century of weather bureau data. The total of 67.03 inches in New York City was 15% over the previous record (58.32 inches) and 60% above normal. The '72 figure so surpasses the record, one can not help but wonder whether there was a cause. During the last year or two, cloud seeding in the midwest has grown from almost zero to a major activity. Is it possible that, as a result of the seeding, instabilities are created in the atmosphere which increase rainfall a thousand miles eastward, perhaps because more moisture is drawn northward from the Gulf of Mexico or because the dynamics of air motion lead to a slow build-up in moisture after the seeding-induced rain? Such explanations are only uneducated conjecture, but they do illustrate the great complexity inherent in evaluation of the weather modification program.

Inadvertent modification

As scientists and engineers learn more about weather modification, the question arises whether the increases in air pollution, particularly near the cities, are sufficient to cause significant changes in the weather in that region. Studies of this possibility in such locations as St. Louis and the northeast are still at a very elementary stage.

Disconcerting data do come from research on La Porte, Indiana, where the steel mills emit large quantities of particulate matter. Comparison with neighboring areas seems to indicate a 30-40% increase in rainfall over the past three decades, and a similar growth in the number of thunderstorms. In the northeast, some environmentalists have suggested that lead particles from auto exhaust may be increasing precipitation (Table 1).

Cloud cover	5-10% more
Winter fog	100% more
Summer fog	30% more
Total precipitation	5-10% more
Relative humidity	2-8% less
Average temperature	0.5-1°C more
Average wind speed	20-30% less

Table 1 Effects of converting a rural area to urban use (from H.E. Landsberg, Climates and Urban Planning, in "Urban Climate," The World Meteorological Organization, Geneva, Switzerland, 1970

Answers to such questions are evasive not only because of the lack of a detailed understanding of weather, but also because the large random variations tend to conceal the small changes representing a steady trend. Concern over the far-reaching significance of inadvertent weather modification is leading to much more intensive scientific research on not only weather, but the global environment.[1]

3. Natural Disasters

The national weather modification program is also directed toward the reduction of the effects of natural disasters.[2] In addition to the hurricane-seeding program discussed at the beginning of this chapter, the following projects are currently underway.

---

[1] SCEP (Study of Critical Environmental Problems), "Man's Impact on the Global Environment," The MIT Press, Cambridge, Mass., 1970

[2] Certain natural disasters are not naturally included within weather modification. For instance, there is a major federal effort to reduce the effects of earthquakes by prediction or anticipation, improving construction techniques and building and highway codes, and educating the public and local agencies in the prevention of earthquake damage.

(1) Cloud seeding is being used in an attempt to reduce hail damage, now running over $300 million per year in this country. The Russians have utilized missiles shot into the threatening clouds and carrying seeding crystals.

(2) Seeding is also used to reduce forest fires caused by lightning (timber losses are $700 million and fire-fighting costs $350 million annually).

(3)Fog at airports and on highways costs $375 million annually in extra transportation expenses, plus the accidents so common on expressways and throughways when fog appears intermittently. Fog dispersal methods are being designed to cut these losses in half. Cold fog dispersal is now achievable, and research is directed toward the more difficult problems with warm fog.

(4) Tornadoes represent a much more difficult problem than hurricanes, but there is early research on tornado modification.[1]

(5) Finally, a major thrust of the total federal program is to vastly improved observation and warning systems. Satellite weather observation is now combined with radar, aircraft, and ground data collection, with the entire system complemented by increasingly complex and accurate computer models for prediction purposes.

4. Conclusion

This first chapter is intended to introduce the general approach we will use as we consider a series of socio-technological problem areas.

---

[1] Edwin Kessler, On Tornadoes and Their Modification, Technology Review, May 1972, pp. 48-55.

As we discuss each topic, the one common characteristic which will emerge is that there usually is no clear solution to the problems. Very often, none of the options is obviously the best. The problems are just too complex and too poorly understood or the interests of different public groups too conflicting to admit a decision which is accepted by everyone as the optimum. Furthermore, as our knowledge and understanding grow, what constitutes the best option may change.

The engineer or scientist, facing this situation, is often dismayed. He is accustomed to working with problems for which there is a correct solution. While the authors are engineering teachers, we also feel strongly that the role of the engineer or scientist in these issues is not to make the decision, but rather to present as completely and honestly as possible his view of the conflicting factors which are relevant. The decision must then be made by the public acting through the normal political and social channels.

The weather-modification program is presented briefly here as a particularly straightforward example in which decisions should be based not only on scientific knowledge and social scientific analysis, but also very decidedly on human considerations.

CHAPTER 2

POPULATION

In the last years of Sukarno's rule in Indonesia, he flatly opposed any program to educate the 100 million residents about family planning or birth control. He fervently believed that if Indonesia was to attain its rightful place in Asian politics, the population would have to grow to 250 million. Consequently, the obligation of the government was to encourage more babies, rather than do anything which might lower the birth rate.

Shortly after Sukarno was overthrown in the mid-1960's, officials of the new government authorized population studies which showed that the population was already destined to reach 250 million. Unless birth control was introduced rapidly, the population would overshoot 250 million by the end of this century. The country would be faced with the severe problems of raising the standard of living and providing employment as well as schools, hospitals, and other public institutions. As a result of these model studies, a nationwide program of education in family planning was introduced.

1. Inertia in the Population Problem

The Indonesian experience illustrates one important characteristic of population problems: there is enormous inertia in population. In other words, even after the brakes are applied, population keeps growing for many years.[1] This is true because, if we have a high birth rate over the past two decades, there are large numbers of girls now less than age 20. Even if the average number of babies each girl will have through her child-bearing years is kept small, the next 30 years will see a large number of births. This in turn will cause many births twenty to thirty years later, and so on.

---

[1] The word "inertia" means two different things. It is often used to mean laziness or sluggishness. Here we are using it in a somewhat different sense as the property of a body or system which tends to keep it moving once motion is underway.

The situation is similar to the problem of stopping a car moving at 60 miles/hour. After the driver slams on the brakes, the car continues to roll forward just because of its inertia. Stopping the car abruptly requires a catastrophic collision with a barrier or an oncoming car. Similarly, stopping population growth abruptly would require such drastic and unacceptable measures as involuntary sterilization of a large fraction of the young people.

This inertia of population growth means that we can not expect immediate changes in the growth rate. If the population is growing now, it will continue in spite of any reasonable action we may take. In other words, there is a long time lag between a new national or public policy and its observed effect. A decrease in the number of babies born per family from today on will not show up as a large drop in population for several decades.

A major goal of this chapter is to study this inertia in more detail. In particular, when we look at a population model for the United States, we will see how this inertial phenomenon appears.

Just because of this inertia, it is important that we be able to predict population trends into the future. Since decades must pass before policy changes have effect, we must anticipate the changes we will want to occur in the future. This is the reason certain alarmists, often anxious to win attention on television, have recently been saying that civilization is doomed, we have already passed the point of no return in population, it is now too late to do anything.

A saner approach was taken by Congress and the President in July, 1969 when they established The Commission on Population Growth and the American Future. In his statement, Nixon said,

"One of the most serious challenges to human destiny in the last third of this century will be the growth of the population. Whether man's response to that challenge will be a cause for pride or for despair in the year 2000 will depend very much on what we do today. If we now begin our work in an appropriate manner, and if we continue to devote a considerable amount of attention and energy to this problem, then mankind will be able to surmount this challenge as it has surmounted so many during the long march of civilization."

Three years later, the Commission report appeared with an extensive study of the characteristics of the population problem in the United States and the options available to the American people.[1]

In this chapter, we will look at population models (particularly for the United States) and then consider the political and social policies which are available to achieve a desired population size in the future. In order to introduce the idea of population models, we first

---

1. The final report was published by the Government Printing Office in the summer of 1972. An advance copy, essentially identical, was published in paperback in March 1972 by New American Library, Inc., New York, N.Y. (W5219) under the title, "Population and the American Future."

consider two simpler cases before turning to the problem of human population.

2. Sea Lamprey in the Great Lakes

The sea lamprey in the Great Lakes have been the basis for one of the earliest and most intensively studied population problems.

The sea lamprey (also called the lamper eel) is a vertebrate which normally lives along the shores of the North Atlantic, in both Europe and North America. Visitors to Puerto Rico can see the eels along the rocky parts of the coast. About three feet long, the lamprey bites its victims and sucks the blood and fluid from the body. After biting a fish, it also secretes an anti-coagulant so that the victim's blood continues to flow. To spawn, the lamprey migrates into fresh water, where the female lays as many as 60,000 eggs and then dies.

A smaller variety of sea lamprey has existed for centuries in Lake Ontario, which is directly connected to the Atlantic Ocean. There have been no lampreys in the other Great Lakes because Niagara Falls (Fig. 2-1) has blocked their travel in from the ocean.

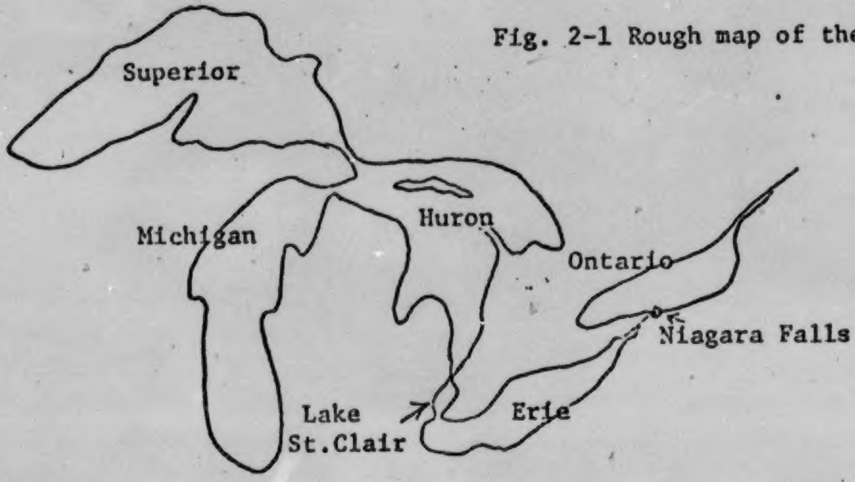


Fig. 2-1 Rough map of the Great Lakes

In 1829 the Welland Canal was completed around Niagara Falls, and ships could move from the Atlantic Ocean into the inland ports. In 1921 (nearly a century later), a sea lamprey was found in Lake Erie. It either swam through the canal or attached itself to the hull of a ship passing through. In 1934 a lamprey was found in Lake St. Clair, in 1936 in Lake Michigan, and in 1945 in Lake Superior. A new species had been introduced into the Great Lakes.

The ecological effects were astonishing. Fishing for lake trout had been a major industry and recreation in the four states on Lake Michigan. The history of the pounds of lake trout caught each year shows the effects of the appearance of sea lamprey:

Year	Millions of pounds
1943	6.9
1944	6.5
1945	5.4
1947	2.4
1949	0.3
1950-3	Essentially 0

In seven years, the lake trout were annihilated. (There was also an effect on the whitefish population. There was a temporary dip, but then actually a slight rise because the trout preyed on the whitefish).

The response of the state and federal governments was typical. Because there had been no anticipation of the problem and no careful collection of yearly data, positive action was not taken until 1948. By then, the data show that the annihilation of trout was well underway and probably unstoppable. The inertia of the system meant that the situation was entirely out of control.

1948 saw the start of a large program to trap the lampreys mechanically. By 1951, the lake trout had disappeared and the lampreys continued to multiply. An electrification program was started. All streams flowing into Lake Superior were equipped with electrified weirs or screens in the hope of electrocuting the lampreys as they swam upstream to spawn or downstream after growing to maturity.

By 1958 it was apparent that the trapping and electrification programs were not successful. The trouble was that the very large number of eggs laid by each female meant that there had to be essentially 100% success in control to prevent population growth. The traps just did not catch enough adults, and the weirs often were broken by debris and ice floating through the streams. Finally, in 1958 it was discovered that some females laid their eggs at the stream and river mouths, rather than upstream. The larvae then lived five years there before transformation to the adult phase, lasting about a year and a half.

In 1958, the final part of the three-pronged program started, with larvicides distributed wide. This chemical "warfare" was aimed at the larvae during the five years between spawning and adulthood.

By the mid-1960's, it was apparent that the sea lamprey could not be wiped out, but the population was at least controlled. By this time, however, the commercial and recreational fishing had been affected disastrously. In addition, swimming on the beaches of Lake Michigan was becoming impossible because of the thousands of dead alewives which repeatedly had to be removed by bulldozers. The alewife population (a small fish) previously had been controlled by the trout predator; now it was free to grow in a friendly environment.

Finally, in the 1960's major programs were launched to stock the lakes with trout from Lake Ontario and coho salmon. The Ontario trout, which had coexisted with lampreys, had developed the ability to avoid the lampreys and also to survive lamprey attacks. The salmon were predators on the alewives and also provided sporting and commercial fishing. By 1970, a reasonably balanced system seemed to exist once again.

The story of the sea lamprey in the Great Lakes may well be unfinished. It is clearly very difficult to anticipate possible future developments. Certainly, however, there is today a vastly greater understanding of the nature of such "population" problems, and the importance of careful data collection so that models can be made which will predict major changes well before they actually occur. The inertia of the system means that corrective measures must be taken very early if they are to have any chance of success.

3. Buffalo

In 1830, there were at least 40 million buffalo (correctly called bison) roaming the western United States. Estimates vary upward depending on which encyclopedia one consults, since no census was taken at the time. In any case, buffalo dominated the landscape the way no other animal has in history. At an average of 1000 pounds each, the buffalo represented 40 billion pounds of bio-mass (compared to only 25 billion pounds of people in all the United States today).

In 1830, the railroad arrived and the rapid westward expansion of the United States began. By 1887, there were only 200 buffalo left.

In this slaughter, animals were often killed for only the tongues and hides. An average of only 20 pounds of meat (of a possible 500) per buffalo was eaten. The peak was reached in 1872 when national heroes like "Buffalo Bill" Cody led the killing of more than seven million.

In less than sixty years, the lack of any sort of policy led to the destruction of what could have been a major source of meat for today's population of this country. A single buffalo could provide the entire meat supply for at least five people for a year.

How could the nation have determined a suitable resource management policy in 1830? In other words, how could a Commissioner of Buffalo have decided how many of the animals should be killed each year in order to keep the total population constant? This is essentially the same problem we face today in setting the rules for the fall deer-hunting season. From a census of the deer population, the state hunting commissioner decides how many can be killed without depleting the total population.

An intelligent policy in 1830 required a knowledge of the way the buffalo population was changing because of natural causes. What are the birth rate and death rate? In other words, what is a model of the population without any harvesting (or killing by human beings)?

Recent studies on buffalo living now in protected areas have indicated the following facts:

1. Buffalo reach maturity at age 2.

2. 90% of the females age 2 or older have one calf a year on the average.

3. 53% of the calves are male, 47% female.

4. 30% of the calves live for two years, or to maturity (infant mortality is high for most animals).

5. 10% of the mature beasts die each year (from tuberculosis, drowning, predation, and so on).

From these data, we can construct a model showing the population of buffalo each year.

Before we develop this model, it is simplest to work first with the population of females. The number of females determine the babies or calves born each year, if we assume only that there is an adequate supply of males. Once we know the number of females, we can find the number of males by a similar model. Consequently, in the remainder of this section (and in the later discussion of human population in the U.S.), we focus exclusively on females.

Now we start constructing the model--just a step-by-step way to find the population each year from the population in the earlier years. We start with 1830, when we assume there are 20 million adult females:

	1830	1831	1832	1833	1834	1835
Adult females	20,000,000					
Female calves almost two yrs.	(a)					
Female calves just born	(b)					

In this table, we need to fill in the entries labelled (a) and (b). (a) is the number of calves who will become adults in 1831. In other words, it is the calves born in 1829 who will reach maturity (age 2) in

1831. Our data on birth rates and infant mortality show that (if we assume 20 million adult females in 1829) there were 18,000,000 calves born then (90% of the adult females). 47% of these were females, or 8,460,000. 30% of these live for two years to adulthood and will become adult females in 1831. Thus, entry (a) is 0.30 x 8,460,000 or 2,540,000.

Entry (b) is the number of female calves born in 1830, or 8,460,000. The table now becomes

	1830	1831	1832
Adult females	20,000,000	(c)	
Female calves about 2	2,540,000	(d)	
Female calves born	8,460,000	(e)	

We now turn to 1831, where we need to find three entries. They are determined as follows:

(c) The number of adult females in 1831 is the sum of two parts:

18,000,000 who survive from 1830 (90% of the adults in 1830)
2,540,000 who become adults in 1831

Thus

(c) is 20,540,000, or 540,000 more than there were in 1830

(d) of the 8,460,000 female calves born in 1830, 30% live through 1831, so this entry is

0.30 x 8,460,000 or 2,540,000

(e) 90% of (c) is the calves born in 1831, 47% of these are female, or entry (e) is 8,650,000

If we continue the table one more year, we obtain

	1830	1831	1832
Adult females	20,000,000	20,540,000	21,030,000
Females almost 2	2,540,000	2,540,000	2,600,000
Female calves born	8,460,000	8,650,000	8,900,000

This table is a population model for the female buffalo. It shows how the number of adult female buffalo grows each year because the number dying is lower than the number reaching adulthood. If we continue the table farther, we find that the population continues to grow indefinitely.

We can now ask the original question with which we started this section: How many adult females can we allow harvested each year to still maintain the population constant? We now need a table with four rows--the last the number to be killed each year. We assume that in 1830, the season starts after the calves are born, so that the number of calves in 1830 is the same as with no harvesting. Then the table takes the following form:

	1830	1831	1832
Adult females	20,000,000		
Females almost 2	2,540,000		
Female calves born	8,460,000		
Females harvested	(f)		

What do we enter in (f)?

If the adult population in 1831 is to be constant at 20,000,000, we can harvest 540,000 in 1830. (With no harvesting, the adult population would grow by 540,000, as shown in the preceding table). Thus, if entry (f) is 540,000, all entries for 1831 are exactly the same as for 1830:

	1830	1831	1832
Adult females	20,000,000	20,000,000	20,000,000
Females almost 2	2,540,000	2,540,000	2,540,000
Female calves born	8,460,000	8,460,000	8,460,000
Females harvested	540,000	540,000	540,000

Each year we can harvest 540,000 adult female buffalo. The adult population stays constant at 20,000,000. If we calculated a similar table for males, we would find we can harvest 860,000 adult males (a larger figure because there are more males born). In total, then, if we can be sure the right proportion of males and females are killed, we can harvest 1,400,000 buffalo a year--enough to provide the entire meat supply for 7 million people.

The most important feature of the above population model is that it permits us to see year-by-year what is happening. We can adjust each year the number of buffalo harvested to obtain control over the population. For example, if there is a drought one year and 20% of the adult population dies, the population dips. As soon as we recognize this decrease, we can reduce the number harvested to allow the population to build up again to 20,000,000 adult females.

In other words, this is what is called a dynamic control model: it shows the year-by-year situation. We can try small changes in harvesting or control policies and measure their effects on the population. In this way, intelligent control is possible even when the birth and death rates are not accurately known, or when natural events such as major weather changes can not be anticipated.

In ecology, the population table developed above is called a resource management model: the quantitative information needed to develop a policy for the management of a natural resource, here the buffalo supply. As we see in the next section, exactly the same kind of model is useful in predicting human population. When we study people, we obviously are not talking about the management of resources, but rather the prediction of future trends so that we can prepare to meet the needs in energy, transportation, housing, education, and the like.

4. U.S. Population

Population predictions for the United States or for the world are a popular pastime in the press and magazines. The fact that the world population is growing has been of interest since the time when Malthus foresaw widespread starvation when population exceeded the numbers that could be fed. A favorite of predictors is to look ahead to the year 2430 (or some such) when people will be standing five deep on all available land on the earth, if the present growth rate continues.[1] (Such a prediction is clearly ridiculous; well before that time there just will not be space available to generate babies).

A much more realistic problem is to try to predict the population of the United States a few decades into the future, rather than centuries.

---

[1] The popularity of such wild predictions is partly the result of widespread public ignorance about population problems. As just one instance, 60% of Americans in a recent survey were unable to state the U.S. population within 50 million; 84% could not give the world population within one half billion. While such data may not be very important, they are indicative of the general lack of knowledge on many aspects of population problems.

Because of the inertia of the system, reasonably accurate predictions should be possible. For instance, most of the mothers from now until the year 2000 are already alive, so we should be able to estimate the number of babies likely to be born during the rest of this century. If there is no major war or catastrophe, the deaths likewise should be reasonably predictable. Let us see how we would develop such a population model for the United States.[1]

Once again, as in the buffalo case, we center our attention on females only. And again we divide the population by ages, since the likelihood of women having children depends on age. If we were making an accurate model, we would use groups covering five years--that is, females 0-4, 5-9, 10-14, and so on. In order to simplify our model, we take larger age groups--every 15 years. Thus, we are working with females 0-14, 15-29, 30-44, and over. From the 1960 census, we obtain the following data:

	1960	1975	1990	2005
0-14	27.4			
15-29	17.7			
30-44	18.4			
45 and over	26.4			

Female population
in millions

---

[1] The model that follows is based on the article by Nathan Keyfitz, "How Crowded Will We Become," on pages 297-309 of the book: J. M. Tanur et al, "Statistics: A Guide to the Unknown," Holden-Day, Inc., San Francisco, Calif., 1972. The paperback book is an exciting and impressive collection of articles indicating many applications of modern statistics to problems of the real world.

We will predict the population every 15 years through the turn of the century.

The first three entries cover 15-year age spans. Clearly, the population 15-29 is surprisingly small. This is the group born from 1930 to 1945--the depression and war years when the birth rate was very low. The group 0-14 were born during the "baby-boom" post-war years from 1945-1960, when this country reached an average of four children per family. Thus, the data for 1960 reflects the changes in public attitudes over the preceding half century, as the desired family size responded to economic and social forces.

Now let us suppose it is 1966 and we wish to try to predict population into the future, perhaps as far as the year 2005. We have detailed data on the birth rates and death rates during the year 1965. In other words, with regard to death rates, we know the small percentage of women of each age who are likely to die in a year. In the same way, we know what fraction of women of each age will have a baby daughter during the year.

From these data, we can calculate how to find the appropriate entries for the 1975 column in the above table--that is, the prediction of the population for the year 1975. Specifically, the number 15-29 in 1975 is 0.9924 of the group 0-14 in 1960. In other words, 0.2 million of the females die during this 15 years as they age from 0-14 to 15-29. (In this table we have omitted the population over 45 to simplify our calculations. Instead of finding the total population, we will focus on the three youngest age groups).

	1960	1975	1990	2005
0-14	27.4			
15-29	17.7	27.2		
30-44	18.4			

Similarly, the 30-44 group in 1975 is 0.9826 of the 17.7 million alive in 1960, or

	1960	1975	1990	2005
0-14	27.4			
15-29	17.7	27.2		
30-44	18.4	17.4		

Thus far, the table or model is derived from 1960 census data and the death rates as measured in 1965. Since these death rates are not likely to change much from 1960 to 1975, the estimates can be expected to be rather good. It is highly likely that there will be close to 27.2 million females 15-29 in 1975.[1]

Now we must determine the population 0-14 in 1975; these are the baby girls born during the 15 years. Again we use data from the year 1965: what fraction of women at each age had baby girls during the year. We assume these same fractions will describe the births from 1960 to 1975. From this information, we find that the number of girls still living and 0-14 in 1975 will be the sum of three parts:

0.4271 x female population 0-14 in 1960

0.8498 x females 15-29 in 1960

0.1273 x females 30-44 in 1960

---

[1] It is also noteworthy that the death rates are quite small. If we were modelling the male population, we could expect appreciably more deaths because of accidents, homicide, suicide, alcoholism, heart attacks, and so forth.

The total of these three numbers gives 29.1 as the number for the 0-14 entry for 1975. In other words, if the birth rates of 1965 truly describe the 1960 to 1975 period, we can anticipate 29.1 female girls under 15 in 1975. (Hence we assume women over 45 have a negligible number of children).

We can then calculate the 1990 column in the same way from the "data" for 1975. From 1990, we can go on to 2005. The complete table then takes the form

	1960	1975	1990	2005
0-14	27.4	29.1	37.7	44.1
15-29	17.7	27.2	28.9	37.5
30-44	18.4	17.4	26.7	28.4

Female population in millions (based on 1965 birth rate)

Several comments should be made about this population model.

(a) The three entries in the 1960 column are census data--in other words, facts.

(b) The death rates are not likely to change very much in the next 30 years. Consequently, the predictions depending only on death rates are likely to be very good. Thus, we have confidence in those entries circled in the table.

(c) The entries not circled depend on birth rates or the patterns of human behavior. In the U.S., a trend to small or large families seems to sweep like an epidemic across the country. If we had used the lower birth rates of 1971 instead of 1965 as a basis for prediction, the table would have the form

	1960	1975	1990	2005
0-14	27.4	25.3	31.4	33.3
15-29	17.7	27.2	25.1	31.2
30-44	18.4	17.4	26.7	24.7

Female population in millions (based on 1971 birth rate)

The circled entries are unchanged; the other entries are reduced. By 2005, this table gives a total population 21 million less than the preceding table (over 40 million less if we include males).

This last comment emphasizes that the model predictions are dependent on human behavior which we can not hope to anticipate. The farther into the future that we try to predict, the more sensitive our model is to these unknown factors. We can predict reasonably well 15 years into the future; we just do not know what will happen 45 years away (even if no major war or depression or new contraceptive introduces an entirely unexpected factor).

5. Use of the Population Model

The population model provides a basis for economic and educational planning. Because many years are required to educate professionals (doctors, teachers, nurses, and so on), the nation should anticipate the needs in order to expand and contract educational programs. Otherwise, we are likely to encounter the situation that in 1972 (for example) there is a surplus of engineers. Freshman enrollment in engineering falls and certain schools close. Five years later, the low output results in a severe shortage, starting salaries soar, and large numbers of freshmen are attracted. Five more years see a serious over-supply. This unstable, oscillatory situation grows more violent each cycle, with large numbers of young people frustrated by the system.[1]

In the particular case of elementary school teachers, our model of the last section shows clearly why the Bureau of Labor Statistics recently predicted a severe surplus by 1980. In the last table (based on the 1971 rate), the number of females 0-14 in 1975 is slightly less than in 1960; even the table based on 1965 births indicates only a climb from 27.4 million in 1960 to 29.1 in 1975. The rapid expansion of elementary schools in the 1960's which resulted from the baby boom of the 1950's will be followed during the 1970's by a static situation at best. Indeed in 1972 elementary school enrollment was only 35.9 million, a half million less than in 1971. A reduction in 500,000 students means about 2,000 fewer teachers. It is not surprising that several states are sharply reducing the number of education programs, and that the newspapers carry frequent stories of education graduates migrating to Australia to find job opportunities.

Both the last two tallies indicate, however, a rise in elementary-school enrollments by 1990 (when the many girls born in the '50's have recently had children). Thus, the model gives a prediction of both the people likely to be entering the job market and the jobs available. Obviously careful economic and governmental planning requires much more detailed models than we have derived in the last section, but our analysis does indicate the principles from which such models can be built.

---

[1] Medical education has consistently avoided this sort of wild oscillation by stringent control over enrollment in spite of the frequent criticism that medical schools are unnecessarily small.

6. Exponential Growth

The model based on 1965 birth rates is reproduced below. If we look at the data in any one row, we find that population grows as shown by the solid curve of Fig. 2-2. There is a rather steady, smooth growth represented by the dashed line; on top of this or added to it is an oscillating part.

	1960	1975	1990	2005
0-14	27.4	29.1	37.7	44.1
15-29	17.7	27.2	28.9	37.5
30-44	18.4	17.4	26.7	28.4

Female population in millions (based on 1965 birth rate)

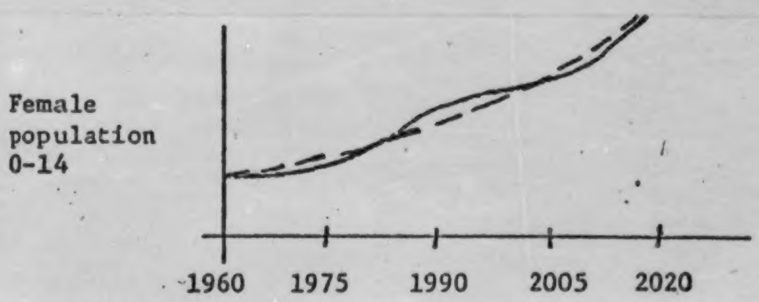


Fig. 2-2 Population growth

The steady-growth part is a long-term tendency of population to grow simply because people like to have more than two children per couple. The oscillating part arises from the short-term changes in public attitudes toward family size--changes caused by war, depression, and the like.

If our prediction is continued into the future, we find that the short-term oscillating part eventually dies out. The prediction follows a simple, smooth growth curve.[1] This smooth curve (the dashed line of Fig. 2-2) indicates the long-term population growth rate.

If we study this long-term growth, we find that it corresponds to a rate of 1.3% per year. In other words, every year the population increases by 1.3%.

The situation is exactly similar to the growth of a savings account in a bank, with the interest deposited at the end of each year (in banking terms, the interest is compounded annually). If the interest rate is 5% and we start with $1000, after one year we have $1050. The next year we receive 5% of this total, or $52.50; hence after two years we have $1102.50. The dollar growth the second year is slightly larger than during the first year.

Each year the growth is 5% of the total at the beginning of that year. If we plot the growth, we find a curve such as Fig. 2-3. The larger the value of the curve, the more the growth that year.

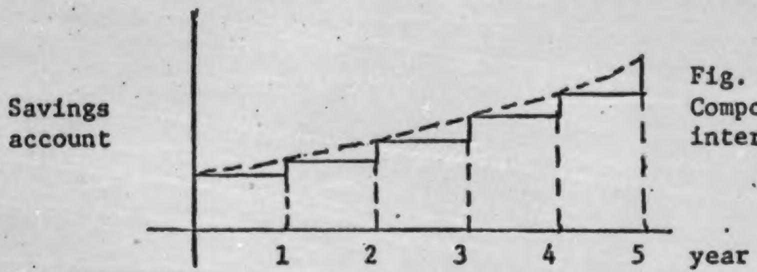


Fig. 2-3 Compound interest

---

[1] Actually, if women had children only at or near age 25, the oscillatory part of the prediction would tend to grow in size. We call such a system "unstable." The model is obviously of no value in predicting population once we are so far in the future that the large oscillations cause violent changes. Fortunately, in the case of human population studies, women have babies over a wide age span. As a result, oscillations caused by the baby boom of the 1950's tend to disappear gradually.

The doubling rule

In the case of the savings account, if we took the 5% interest home each year and kept it "safely" under the mattress, we would have $1000 in the bank and receive $50 each year. In 20 years, we would have $1000 under the mattress; our money would have doubled.

If we invest the interest every year (that is, we obtain compound interest), each year the growth is slightly larger than the year before. Now how many years must pass before the money in the bank doubles? If we do an exact calculation using algebra, we find the answer is 14.2 years. Actually there is a simple rule which works for all compound-interest problems (or population growth problems). It states that

> The number of years to double is about 72 divided by the percent growth per year.

In growth problems, this rule is a very important and useful tool. For example, if the interest rate is 5%, we must wait about 72/5 or 14.4 years for the total to double. (We said above the correct answer is 14.2 years; this difference demonstrates the typical accuracy of the rule).

If the U.S. population is growing by 1.3% each year, we can expect it to double in 72/1.3 or 55 years.[1] A population of 205 million in 1970 will become 410 million by the year 2025.

---

[1] The correct answer is 54 years. Again our rule works, since no one is going to worry if we are one year off. The rule works well as long as the interest rate is less than 20%.

Another form of the doubling rule

The rule can also be stated in another useful way:

> The percent increase per year is about 72 divided by the number of years to double.

If we know that a certain growth situation results in some quantity doubling in 18 years, we can immediately state that the annual growth rate is 72/18 or 4%.

In both forms of the doubling rule, we need not measure time in years; we can use any convenient length of time. For instance, if a baby weighs twice its birth weight at nine weeks, the average weekly increase is then 72/9 or 8%. We could also state: the weight doubles in 63 days; hence the growth rate is 72/63 or 1.1% per day.

Thus, the two doubling rules allow us to slide easily back and forth between the two descriptions of growth: the time to double and the percent increase each year or other unit of time.

Exponential growth

The particular type of growth considered in the above examples is called exponential growth. This term means that the quantity or size grows by the same multiplying factor in each unit of time.

This is an extremely important concept in many different system studies. In 1970 in Boston, one in every five people were on welfare. With a population of 700,000, this means 140,000 on welfare. If the growth is exponential for the number on welfare and the rate is 6% per year, 1971 should see an increase of 8,400 (to a total of 148,400). In 1972, the increase is 8900, or 500 more than in 1971. When the growth is exponential, the problem is not only continuously worse, but also the amount by which it grows continually increases. In general terms, the rate at which the situation worsens grows directly as the severity of the problem.

In popular speech, we often describe such a situation as runaway. Runaway inflation menas the dollars needed to buy a house, for example, grows exponentially. In Brazil a few years ago the inflationary rate was 20% per year. In less than four years, the cost of a house doubled.

When growth is exponential, the danger is that the problem may get away from us before we realize its seriousness. There is a classic example of this in ecology. A man owns an estate on which a pond is located. He finds that the water surface is gradually being covered by an odorous plant growth. He notices that a square foot is covered the first day, two square feet the second day, and four square feet the third day. He decides he will take corrective action when half the surface is covered; by that time the problem will be serious.

Unfortunately, by that time it is too late. The growth doubles each day. Consequently, when he moves into action, he has only one day to correct the situation.

Sigmoid growth

Actually exponential growth usually characterizes the early portions of a phenomenon. The annual sales of the Xerox Corporation can grow exponentially during the company's early years, for example at 18% per year or doubling every four years. Obviously such growth can not continue indefinitely; the Xerox sales can not exceed the U.S. Gross National Product, and the slowing of the growth rate will occur well before that time.

An actual growth curve tends to have the form of Fig. 2-4, known as a sigmoid curve (sigmoid means shaped like the Greek letter sigma,

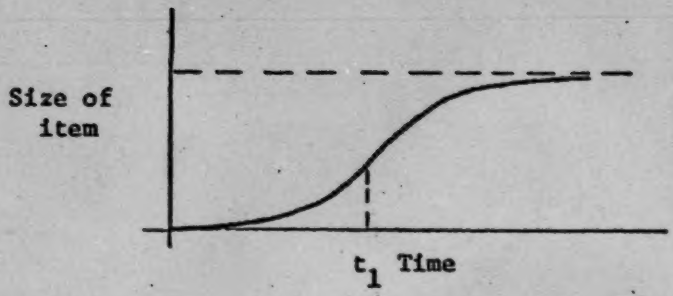


Fig. 2-4 Sigmoid growth

corresponding to s). The growth is exponential up to the time t_1, then the rate of increase starts to slow, and finally the population or size approaches a constant value (or a slowly increasing value).

In any particular growth study, a crucial problem is deciding where we are with respect to t_1. Is the growth going to continue exponentially, or will the growth shortly slow down? In most cases, the answer is not at all apparent. For instance, in planning for new schools in a suburban town, the School Board has to anticipate when population growth will slow. Obviously, if land availability is severely limited, immigration will fall. But usually the immigration decrease precedes this time, since as land becomes slightly scarcer, prices rise and the area is less attractive because of crowding. The difficulty of the problem is proven by the suburbs where elementary schools have been over-built.

7. Graphs of Exponential Growth

The United States is by far history's most efficient producer of solid waste. With only 5% of the world's population, this country generates well over 50% of the total rubbish. The data from the past 45 years are awesome!

1926	50 million tons/year
1940	75
1950	100
1962	150

We desire to plot these data in order to predict the problem magnitude likely by the year 2000 and also by 2050.

First we might ask: Is this exponential growth? This is equivalent to the question: Do the given data represent a constant percentage increase per year? The numbers above do seem to be close to exponential growth. The total doubles from 50 to 100 in 24 years; later it doubles from 75 to 150 in 22 years. It multiplies by 1.5 in 14 years (1926-1940), then in 12 years (1950-1962). While the growth rate may be increasing slightly, probably exponential growth is a reasonable assumption. There is no point in worrying too much about small errors, since the above data are certainly just rough estimates.

We want to plot a graph showing the annual solid waste as it varies with the year. We first select our scales (Fig. 2-5). Since we want to predict to the year 2050, the horizontal scale is chosen to carry the graph

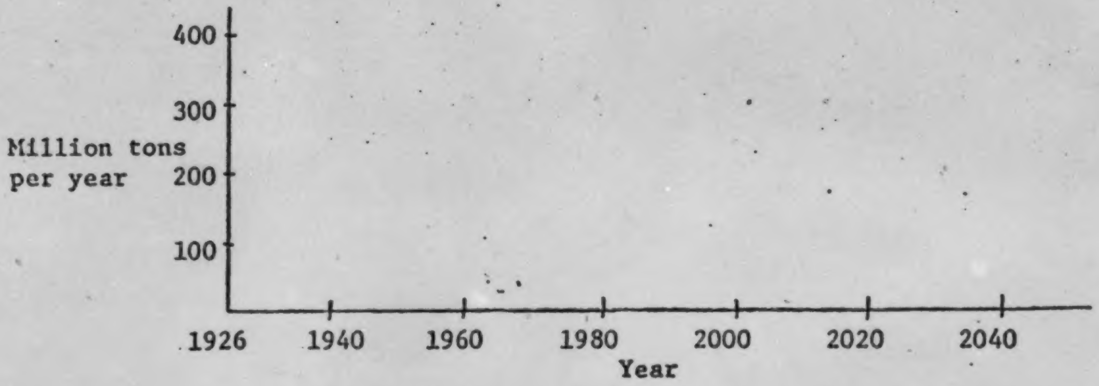


Fig. 2-5 Scales for the graph

this far. The smallest value is 50 in 1926; on the vertical scale we represent this by a small distance.

Now we are ready to plot. The given data are first entered, and a solid line is drawn approximately through these points (Fig. 2-6). If the doubling time is 22 years, prediction points can be determined:

1984	(1962 + 22 years)	300 million tons/year
2006		600
2028		1200
2050		2400

A dashed line in Fig. 2-6 is drawn to indicate the prediction into the future.

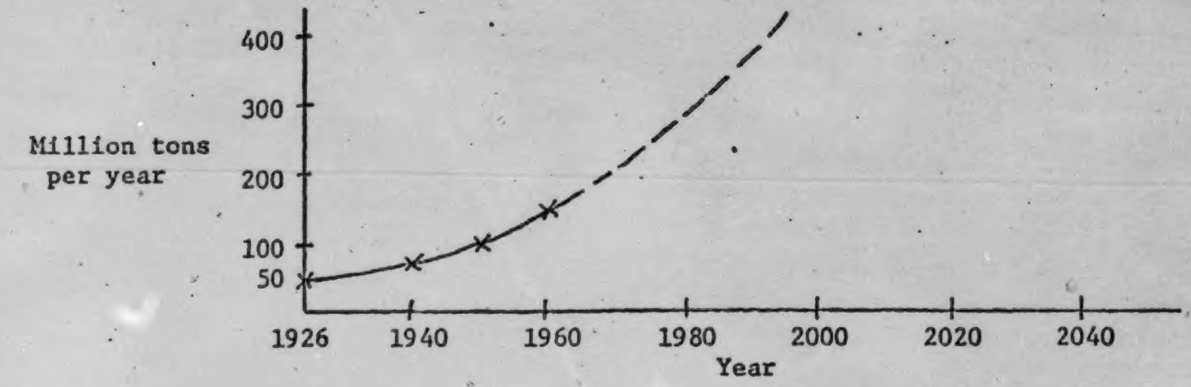


Fig. 2-6 Solid waste prediction

Unfortunately, before the year 2000 the graph goes off the scale. The exponential growth is so rapid, we would actually need a graph nine inches high to show the value by 2050. In order to bring the value at 2050 onto the paper, we would have to contract the vertical scale so much that the curve from 1926 to 1950 would almost lie on the zero line.

These graphing difficulties always arise with exponential growth. They can be avoided by using a distorted or nonlinear vertical scale. Instead of each division vertically corresponding to a fixed number of tons/year,

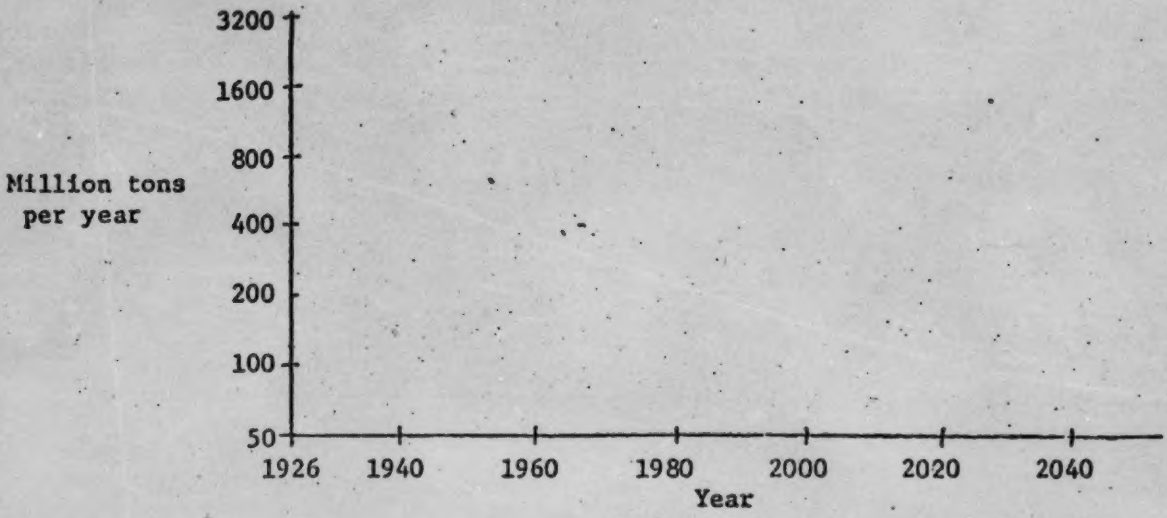


Fig. 2-7 Vertical scale for an improved plot

we make each division represent the same factor of growth. In other words, we label the vertical scale as shown in Fig. 2-7. The minimum value of interest is 50, so we choose this at the bottom. One division up corresponds to a doubling to 100; a second division doubles again to 200, and so on. Each vertical division need not represent doubling; multiplication by any constant factor can be used.

Once the scales are fixed, the given data points can be inserted. With this type of vertical scale, the graph is a straight line. This is just what is meant by exponential growth. The years horizontally required for the curve to move upward one division are the same everywhere since each upward division corresponds to multiplication by the same factor.

Thus, exponential growth is represented by a straight line when we use this distorted vertical scale, called a "logarithmic scale." Such a logarithmic plot allows us to answer the question originally posed: Do the given data correspond to exponential growth? If these data fall on a straight line on such a plot, they do indeed represent exponential growth.

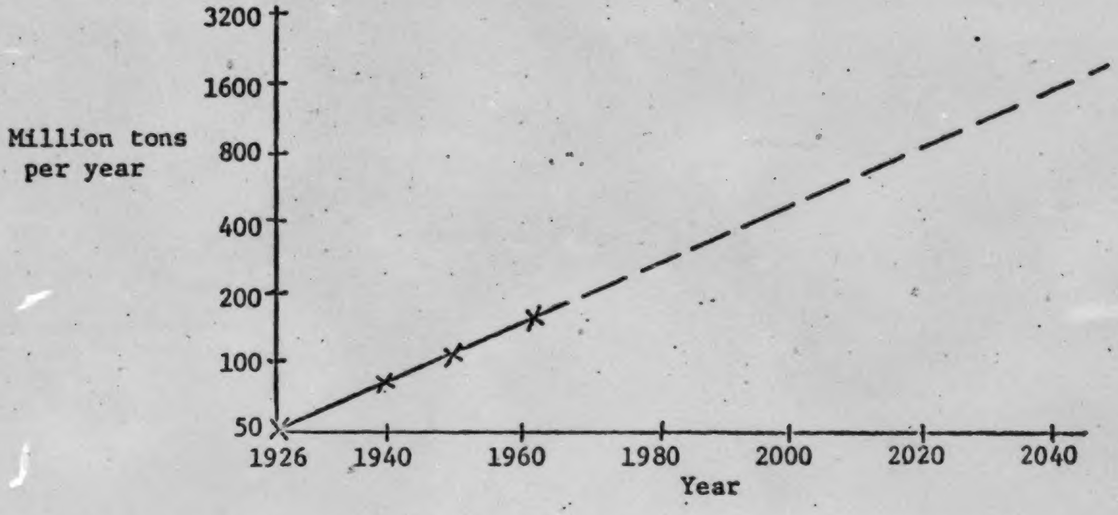


Fig. 2-8 Straight-line plot of solid waste generated (logarithmic vertical scale)

Furthermore, with this type of plot, prediction is trivial. We need only extend the solid line with a straight edge (Fig. 2-8). The figure shows that the prediction gives about 2500 million tons/year in 2050, 500 million tons/year in 2000.[1]

Graphs of the type described above (i.e., logarithmic plots) appear frequently in newspapers, magazines, and other media. The two advantages are that one graph can show an enormous range of variation of the quantity being plotted, and that a straight line represents the exponential growth which is so very common in describing social, environmental, and economic signals.

---

[1] Such graphs and predictions are normally not used for precision work, since the data just are not that good. We should point out that the vertical scale is distorted. Thus, 800 is one division above 400 and is 2 times 400. One half division up would be $\sqrt{2}$ times 400, not halfway between 400 and 800. If one vertical division corresponded to multiplying by 3, one half division would mean multiplication by $\sqrt{3}$ or 1.7.

8. Possibility of Population Control

In 1972, the number of children per family in the United States dropped below 2.1 -- the level needed to establish a constant population on a long-term basis. In spite of this, the inertia of the system means that the population will continue to increase at an average rate of 0.7% per year until the end of the century, and 300 million population will be reached by 2021. After this, the levelling off will occur.

If social, economic, and attitudinal changes occur after 1973 and we return to a norm of 3 children per family, the population will reach 300 million by 1996, and 400 million by 2014. Thus, any rational estimate of future population to the end of this century simply has to be based on a guess as to the reproductive habits of the people. Here past history (Fig. 2-9)is not very comforting, as the reproductive rate has varied from a high of 4 CPF in the late 1950's to 2.1 in 1972. There is no logical basis to estimate with confidence what this rate will do over the coming decade.

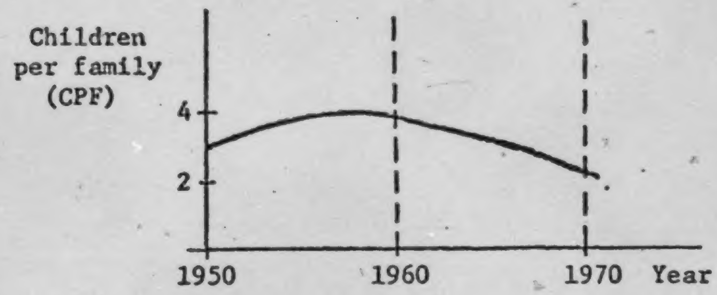


Fig. 2-9 Children per U.S. family over last two decades

Regardless of the changes in the next decade in the CPF, the inertia will cause continued population growth as noted above. To achieve zero population growth immediately, we would have to have one CPF this entire generation, 2 CPF in the next generation, and then 3 CPF in the next (to compensate for the low birth rate in the current generation). Even the most ardent advocates of zero population growth can not seriously argue for the federal and social constraints which would be necessary to ensure such an odd reproductive pattern.

Consequently, this country must plan to adjust for at least several decades to population growth. During the next two decades, we can anticipate

	1970's	1980's
New households/year	1.5M	1.5M → 1.2M
Increase in labor force/year	3.5M	2.8M
Percent living in cities	60% → 70%	70% → 78%

Because of such impacts, a national population policy must be developed.

If there is a national decision to attempt to limit population growth, the question that immediately arises is: How can control be achieved? What are the alternatives?

In order to answer this question, the Commission on Population Growth and the American Future attempted to look in depth at the various factors which contribute to growth. From such a study, a relatively small number of tolerable alternatives appear, which we discuss in the remainder of this section.

Immigration

Does immigration contribute significantly to U.S. population growth? Yes, even under the restrictive laws now in effect.

More than 1000 people per day is the net flow into the United States, or approximately 400,000 per year. This compares to 10,000 babies born daily, and 5000 deaths. Hence, immigration accounts for 20% of the annual population growth in the early 1970's.[1] Two thirds of the immigrants settle in only six states (New York, California, New Jersey, Illinois, Texas, and Massachusetts), so there is a particularly acute problem in absorbing the immigrants into the socio-economic system in a few cities. The impact on population growth is particularly pronounced because so many of the immigrants are young adults just entering the child-bearing age.

The problem of controlling immigration has been difficult in the United States because so much of the population growth during the last 150 years has resulted from the open-door policy. The first restriction was imposed in 1882 when Chinese were excluded. About 1900, general limitations were imposed on those with criminal or bad health records. 1924 saw the first general restriction with only two categories allowed to immigrate: (1) Close relations of U.S. citizens and refugees; (2) 150,000 others each year, selected by quota to maintain the existing ethnic distributions in this country.

These regulations to hold the "status quo" lasted until 1965, when a major revision of the immigration laws removed emphasis on national origin and based future decisions on:

(1) Family reunification,

(2) Political asylum,

(3) Needed skills and professions.

This last category has resulted in the heavy in-flow of physicians, attracted by the high salaries (average $42,000) and exceptional clinical and research facilities. (There are 7000 Filipino doctors in this country now, compared to only 6000 Blacks).

---

[1] This 400,000 annually is a net figure. About 37,000 people per year emigrate from the United States, primarily to Canada, Israel, and Australia.

An entirely different category encompasses the illegal immigrants. Nobody really knows how many of these there are. In 1971, 420,000 such aliens were located and deported, in contrast to only 71,000 in 1960, so presumably the in-flow rate is of the same order of magnitude as the legal immigrants. One estimate is that as many as two million aliens may now be living illegally in this country.

Eighty percent of the aliens deported in 1971 were from Mexico. In most cases, these were itinerant farm workers, willing to work for below-minimum wages. In recent years, an increasing percentage of the illegal immigrants have been found in the larger cities, in contrast to earlier years when most stayed near the point of entry.

Thus the immigration, both legal and illegal, constitutes 40% of the annual population growth of the United States. If immigration were shut off, the United States would now be well below the reproductive rate for long-term, zero population growth. In the light of this situation, some proponents of population control argue for both a significant cut-back in legal immigration and rigid law enforcement to decrease illegal entry. They emphasize the unemployment problems already existing in this country and the urban congestion.

To refute these arguments, people in favor of immigration point out that the immigrants in the past have made extremely important contributions to the development of this country. Further, for humanitarian reasons, the United States should continue to be an asylum for political refugees, and a land where opportunity exists without regard to the individual's background.

Unwanted babies

If a reduced rate of population growth is a national goal, one might ask if decreasing the number of unwanted babies could be significant. Here the term "unwanted baby" refers to children not desired by the parents before conception. According to the 1965 National Fertility Study, 20% of all babies fall into the unwanted category. One third of the married couples studied had at least one unwanted child. In terms of the number of children the parents already had

5% of the first children were unwanted
8% of the second
21% of the third
32% of the fourth
45% of the fifth
50% of the sixth and higher

The 1970 National Fertility Study showed similar results, with the clear indication that if unwanted births were excluded, the United States would be well below the reproductive rate corresponding to an ultimate, zero population growth.

In urging a national effort to reduce the number of unwanted births, the majority report of the Commission on Population Growth and the American Future argued for a variety of actions. The passage of the 1970 Family Planning Service and Population Research Act was considered only a first step. The Commission stated that much more research is needed on contraception, and educational programs must be vastly expanded to make family planning services more widely available. Furthermore, many of the state laws on marriage , divorce, abortion, and contraception should be changed. The arguments are made that advertising of contraceptive devices should be allowed, and that physicians should be legally able to give advice to minors without parental consent.

Other factors

Population growth can also be slowed by increased opportunities for women to work. In much of Eastern Europe, where women have been heavily employed since World War II, the birth rate has dropped so low that there is widespread concern. In several countries, national policy is to encourage additional children through all possible incentives and propaganda.

The birth rate can be influenced somewhat through economic incentives and tax laws, although the system is generally inelastic -- that is, changes in the federal income tax rates, for example, are unlikely to have any real effect.

To what extent propaganda is effective is not clear. The sharp drop in the CPF (children per family) from 1959 to 1972 is generally attributed to a variety of causes: the birth control pill, expansion of family planning services, easing of abortion laws in several states, women's liberation and expanded job opportunities, economic conditions such as the rise in unemployment about 1970, and publicity about the population and environmental problems. There seems to be no agreement on the relative significance of each of these factors. Even in the case of abortion-law change, the fact that 400,000 abortions were performed last year in states where it was permitted at the will of the mother is still difficult to interpret. How many of these abortions would have been done illegally? How many of the pregnancies would not have occurred with stricter abortion laws? All that really is known is that the number of unwanted births is certainly being reduced sharply if the shortage of babies available to adoption agencies is a valid indicator.

9. Final Comments

This chapter really only touches the surface of the subject of population models. We owe an apology to the experts in demography and population modelling for the superficial treatment of an exceedingly difficult topic. Perhaps even this brief coverage does indicate the complex nature of population dynamics, with the importance not only of the age distribution of the current population, but also of public attitudes and governmental policies.

The difficulty would be vastly greater if we included consideration of the geographical distribution of the U.S. population. Anyone who has travelled throughout the country recognizes that there is no shortage of land. The trouble is that most of the population lives in the urban areas. In 1900, 60% of the people lived on farms or in villages; by 1970, only 26% were rural and 69% lived in urban areas (cities of at least 50,000 or areas economically integrated with cities). The percentage was even higher in California (93%) and the Northeast (80%). The 29 cities today of population more than one million will grow in number to 50 by the year 2000, and there will be 12 metropolitan areas the size of New York today.

The trouble with such urbanization arises both because of the concurrent intensification of basic problems (environmental, transportation, crime, and so on) and because the city is not where people often want to live. If people are asked for their choice of living environment, they respond as follows:

	Percent living there now	Percent preferring this
Open country	12	34
Small city	33	30
Medium city	28	22
Large city	27	14

While the aversion to the large city seems clear, it is not clear whether urban population density is sufficiently high to cause serious effects on the inhabitants.[1] Experiments with animals, starting with the rats studied by Calhoun in 1962, have shown that serious overcrowding causes increased mortality among the young, lower fertility, neglect of the young by mothers, aggressiveness, somnambulism, sexual aberrations, and other psychotic behavior. To what extent can these results be carried over to a study of the modern city?[2] Since it is nearly impossible to do carefully designed, extensive social experiments on people, there is no answer to this question. Perhaps we can at least be comforted by the fact that there is no clear evidence that today's cities are having similar effects on the residents.

---

[1] One pictures New York City as crowded. Actually the population density of 80,000 per square mile is far below ancient Rome's figure of 225,000, even though there were no high rise buildings at the time of the Caesars.

[2] An interesting article with a bibliography is that by O.R. Galle et al, Population Density and Pathology: What are the Relations for Man, Science, April 7, 1972, pp. 23-30.

CHAPTER 3

DELIVERY OF HEALTH SERVICES

The year is 2040; your granddaughter has just celebrated her 30th birthday at an all-night party. She faces a bleak prospect. There is a real shortage of marriageable young men. Indeed, in the U.S. as a whole, there are 25% more women than men, with the difference almost entirely over the age of 20. Both candidates for President are supporting legislation allowing only young males to immigrate.

Fig. 3-1 shows what has happened. As recently as 1910, there were 6% more men than women. Then the ratio started falling dramatically. In the 1970's, health experts became alarmed and studied the causes. They found that

(1) Young men (under 25) were dying at a rapidly increasing rate from auto accidents, suicide, and murder. Auto accidents caused half the deaths in the ages 15-19.

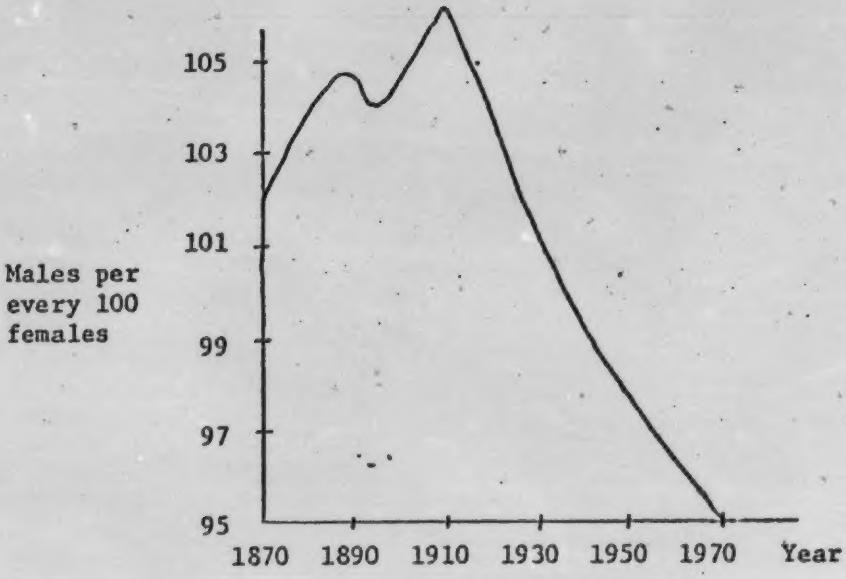


Fig. 3-1 Ratio of men to women in the United States

(2) Old men were dying earlier and earlier from lung cancer, liver disease (related partly to alcoholism), and heart attacks.

As an indication of the problem, the death rates among people 25 through 29 were

White women	71/100,000
White men	167/100,000
Non-white women	201/100,000
Non-white men	455/100,000

Even though these data were widely publicized during the 1970's and 1980's, there was no real demand for public action to change the trends. More and more women were following careers, and the lack of marriage was not considered serious. Then in the early years of the 21st century, the counter-reaction to women's liberation set in, and the common goal of young women was devoting full energy to the raising of a family. Suddenly, the shortage of men became painful. Safe driving laws were strictly enforced, prohibition became the law of the land, smoking was banned in all public places, and severe tax penalties were imposed on any man over 30 who was unmarried for more than one year.

Unfortunately, in the year 2040 these measures have not been in effect long enough to improve the situation. As so often happens, public awareness of a health problem followed by many years the clear signs of the problem. Public action didn't have any obvious effect until decades later. During this long time delay, the problem became serious.

1. Health Services

The high male death rate is just one indication of the current health problem in the United States.

Every U.S. President since Truman has affirmed the right of every American to high-quality health and medical services. With the financial and educational resources available in this country, we might logically guess that the health of the average American matches or exceeds that found in any other country. Yet the few quantitative measures available for the comparison of different countries usually place the United States well down in the list of the advanced nations in the world.

What are the possible reasons for the relatively poor health status of Americans? What positive steps might be taken to improve the situation? What blocks or impedes such action?

The familiar lament we hear nowadays is: If we can send a man to the moon within a decade, why can't we demand the same success when we focus equal energy on improving the health of the people (or indeed on any other problem)? There are, as we shall see, two parts to the answer:

(1) The health problem is much more complex than the moon program. In the health problem, we often do not know whether an experiment is successful or not. There is no simple way to measure how healthy people are.

(2) In this country we have found it exceedingly difficult to focus effectively our national resources on a particular socio-technological problem. In any national problem such as health, attempts to use technology and to do intelligent planning run into political, social, and economic roadblocks. Also, there may be so many alternatives for possible action that it is impossible to decide what to do first--with the result, either nothing is done or the total effort is so spread out, no one area is funded enough to show real results.

In the following pages, we want to look at the general nature of the health problem, and then look at a few possible courses of governmental action (with the emphasis on the problems impeding progress). From this "case study" there hopefully will emerge an appreciation of some of the alternatives both available today in the United States and politically feasible with adequate public understanding and support.

2. The Health Services System

What do we include within the health services system? The physical and mental health of Americans is affected by almost every system or institution. Professional sports hopefully contribute to mental health, urban transportation perhaps causes mental strain, and solid waste disposal influences the urban rat population which in turn affects health.

In order to work with a manageable system, we will arbitrarily exclude mental health and all environmental problems (such as air pollution) where the relation to health is not obvious. In particular, we include:

Preventive health care
Hospital care
Care of the chronically ill or the physically handicapped
Handling of diseases and epidemics
Nutrition

Another aspect, emergency medical service, is considered in Chapter 4.

How large is the system for the delivery of health services in the United States? Our expenditures now total about 75 billion dollars per year, and the total is growing appreciably faster than the GNP. Health is either our largest industry or close to it (this depends on what we call an industry).

During the decade of the 1960's, the change in certain costs in the U.S. were as shown in Fig. 3-2; of the items depicted, the only one to rise more than hospital costs was the price of food items popular with

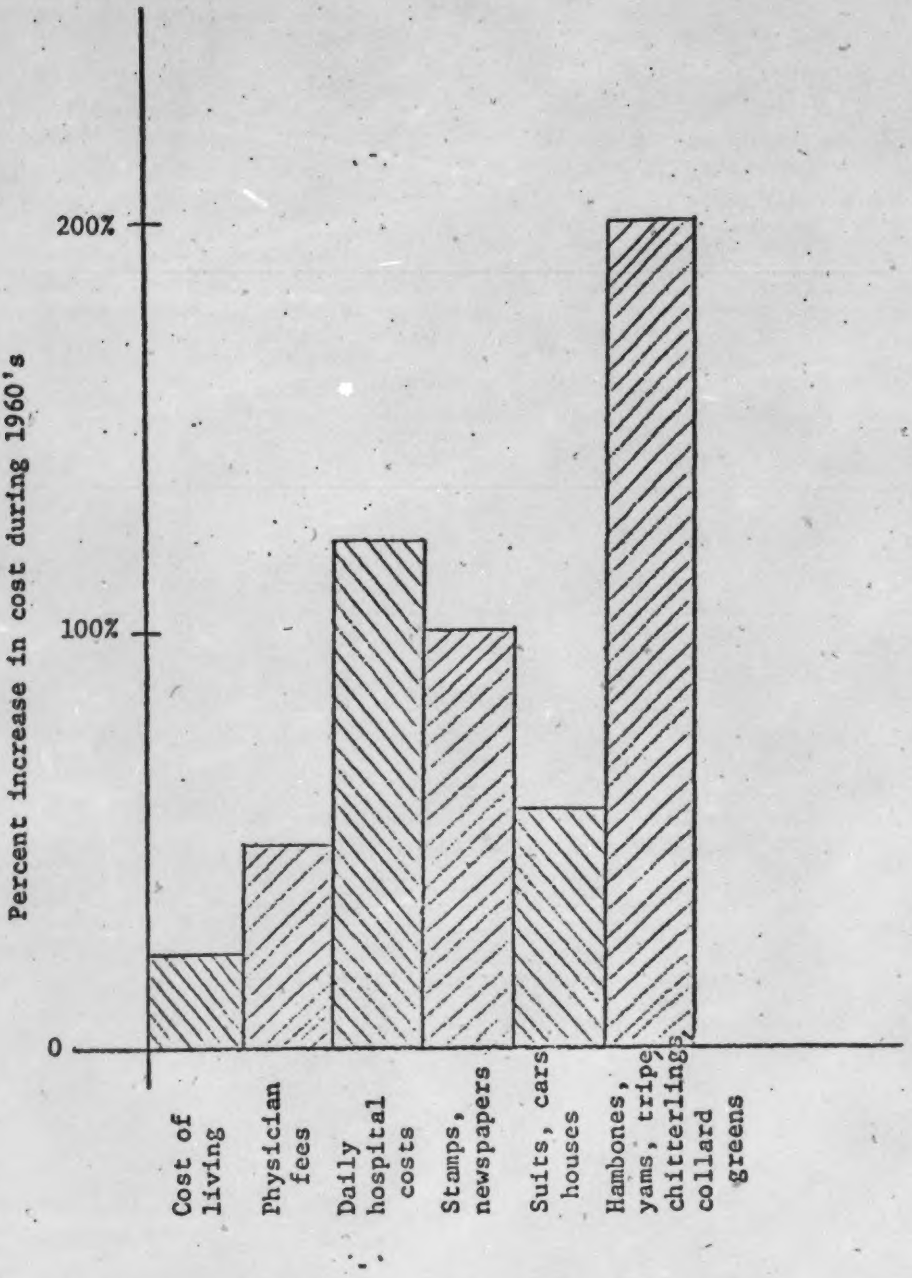


Fig. 3-2 Rises in costs in U.S. during the decade of 1960's

the poor within our population. The phenomenal rise in hospital costs was the result of several parallel developments. First, wages rose sharply, particularly as many hospital workers were brought under minimum-wage laws for the first time. Rapid growth in various insurance programs increased administrative paper work and costly delays in hospitals being paid. New technology, including computers and intensive care facilities, was expensive to purchase and, particularly, to maintain and operate. In this respect, there was very little regional planning or inter-hospital sharing of elaborate facilities; the staff of every hospital wanted the capability for open-heart surgery, even though most of the resulting facilities were grossly under-utilized.

Of the total expenditures for health, 70% come from private sources, only 30% from governmental programs--in sharp contrast to many of the advanced countries of the world where the complete health care system is nationalized.

3. Current Performance of the U.S. System

These expenditures of $375 per person per year in the U.S. are very much higher than other developed countries. We might anticipate a health level significantly better than the other advanced nations of the world. The data concerning the health of the public are not very detailed or complete; for example, we really don't know what percentage of our population between the ages of 40 and 50 are reasonably healthy most of the time. We can estimate the prevalence of certain diseases or body malfunctions, but these data are often subject to error and, even taken together, do not give a good picture of health. Even in such a seemingly simple aspect as the legally blind, estimates of the population range from less than a million to nearly three million.

The lack of trustworthy data appears again and again in public discussions. Within the last two years, there have been repeated arguments in Congress and in the news media about the magnitude of the hunger problem in this country -- arguments made even more confusing by unsubstantiated public statements and occasional willful distortion by news media.

There are, however, some data which are carefully defined throughout the world and which can be used to compare the U.S. with other countries.

Life expectancy

The individual born in the United States in the year 1900 could expect to live 49.2 years; the child born in 1965 estimated his life expectancy as 70.2 years. Figure 3-3 shows how this change of more than 20 years has occurred, and Table 1 gives the detailed data since 1930. The advantage of seven years for women compared to men, plus the fact that the wife is typically three years younger than her husband, explains the very large number of older widows in this country. The typical wife can expect to outlive her husband by ten years. Clearly, we should change our customs to urge young women to delay their marriage until their late 20's and then to marry men in their early 20's.

Two aspects of the life expectancy data are interesting:

(1) There has been a sharp levelling off of the growth in the 20 years since the early 1950's. In 1954, the average was 69.4; in 1965 only 70.2. In spite of the widely publicized medical advances (artificial kidney machines and heart transplants) and the stream of new drugs, life expectancy has not increased appreciably.

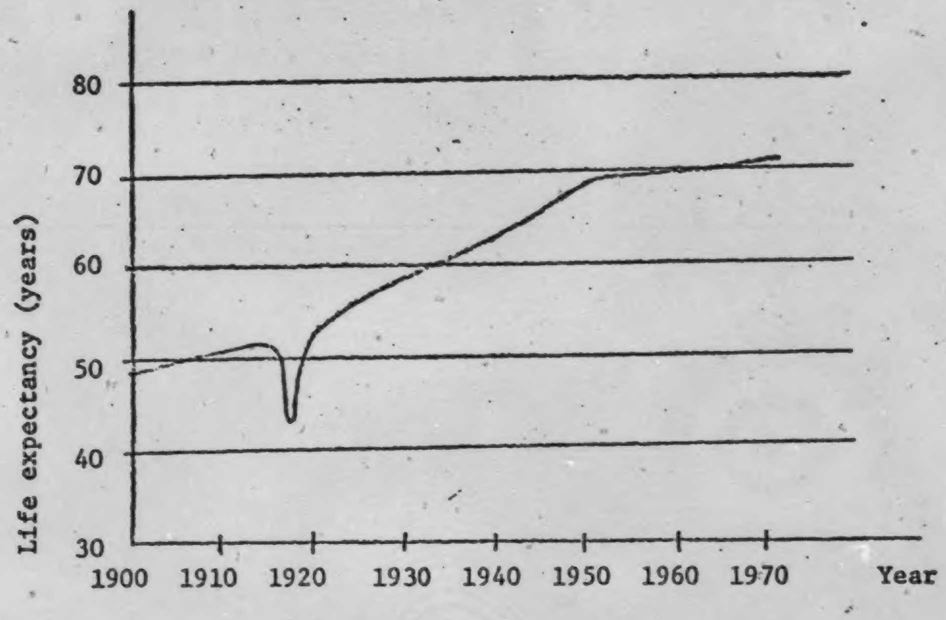


Fig. 2-3 United States life expectancy at birth.

		White		Nonwhite	
Year	Total	Male	Female	Male	Female
1930	59.7	59.7	63.5	49.3	49.2
1935	61.7	61.0	65.0	51.3	55.2
1940	62.9	62.1	66.6	51.5	54.9
1945	65.9	64.4	69.5	56.1	59.6
1950	68.2	66.5	72.2	59.1	62.9
1955	69.5	67.3	73.6	61.2	65.9
1960	69.7	67.4	74.1	61.1	66.3
1965	70.2	67.6	74.7	61.1	67.4

Table 1 Life expectancy in United States

(2) During these two decades, the U.S. has dropped steadily in the world ranking of nations (Fig. 3-4). A complete listing of nations shows that, according to the most recent data available from each country, the United States ranks 13th in female life expectancy, 31st in male life expectancy. Indeed in the latter list, this country ranks below Greece and Yugoslavia, where we usually picture significant sanitation problems. Table 2 is a startling list of areas outranking the United States.[1]

Actually, life expectancy is a somewhat poorly calculated figure and, as we mentioned above, only a rather crude indicator of health. Even the data have to be interpreted with care. Life expectancy is determined each year from the death rates in each age group during that year. For example, for a baby born in 1971, we determine the percentage of 75 year-old men who died during 1971; this is the probability of the baby dying in his 75th year if he reaches 75. We do the same thing for each year (74, 73, 72, and so on down to zero). From these probabilities, we can calculate his life expectancy, although the calculation is not easy.

Thus, the life expectancy of a baby born this year depends on how the older people are dying this year -- rather than on a prediction of how the older people will be dying when he is one of them. Consequently, when there is a major influenza epidemic (1919 in Fig. 3-3), with many deaths among abnormally young people, the life expectancy drops sharply. These data are not published until several years later and hence are of interest only to people who have already survived the epidemic; consequently, the calculations are not too useful.

---

1. We are, of course, comparing the U.S. primarily to other developed nations. If we include the entire world's population, this country looks good. In Gabon, Africa, for instance, male life expectancy is 25 years; in India less than 42.

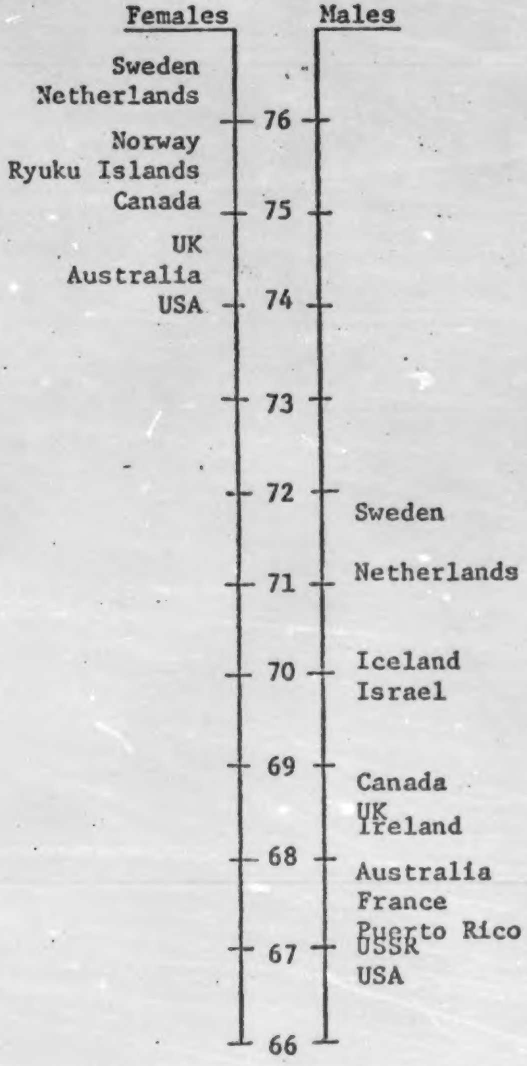


Fig. 3-4 Life expectancy around the world (not all countries are included within the range shown)

Sweden	71.85
Norway	71.03
Netherlands	71
Iceland	70.8
Denmark	70.7
Israel	69.55
Japan	69.05
Ryuku Islands	68.91
Bulgaria	68.81
Canada	68.75
Switzerland	68.72
East Germany	68.72
UK	68.5
Malta	68.45
New Zealand	68.44
Ireland	68.13
Australia	67.92
Yugoslavia	67.74
Belgium	67.73
France	67.6
West Germany	67.55
Greece	67.46
Czechoslavakia	67.33
Spain	67.32
Italy	67.24
Puerto Rico	67.14
USSR	67
Hungary	67
Poland	66.85
Hong Kong	66.74
USA	66.6

Table 2 1971 Ranking of countries by male life expectancy (based on latest available information). These data are taken from the "United Nations Statistical Yearbook, 1971."

Infant mortality

A second measure of a country's health is the infant mortality -- the fraction of babies who are born alive but who die during their first year. This is in some ways a more reliable index than life expectancy, since in most advanced nations accurate and up-to-date data are available. Since 1930, infant mortality in this country has dropped from 65 deaths/1000 live births to less than 20, with the sharpest decreases during the 1930's and 1940's (Fig. 3-5).

In spite of this marked improvement, during this period the United States has fallen far down in the ranking of the nations with which we would expect to be compared (Fig. 3-6).

It is particularly noteworthy that in the U.S. the probability of an infant dying during the first year is almost twice that in Norway or Sweden. (Sweden is the nation often taken as the model for evaluating health delivery because of the generally high ranking in all indices, and the elaborate, nationalized system of health care).[1]

4. Reasons for Poor U.S. Performance

For a nation which spends the most money per capita on health care and in which the objective of the best possible health care for every citizen is a non-political, national goal, these measures of U.S. performance are stunningly disappointing. If these data do give a reasonably accurate indication of our health status, what reasons can be given?

---

[1] We should emphasize that we are comparing the U.S. with other developed nations, not with the world as a whole. In Pakistan, for example, the infant mortality is 142/1000 (more than one in seven babies die during their first year), and in Burma the rate is over 200/1000.

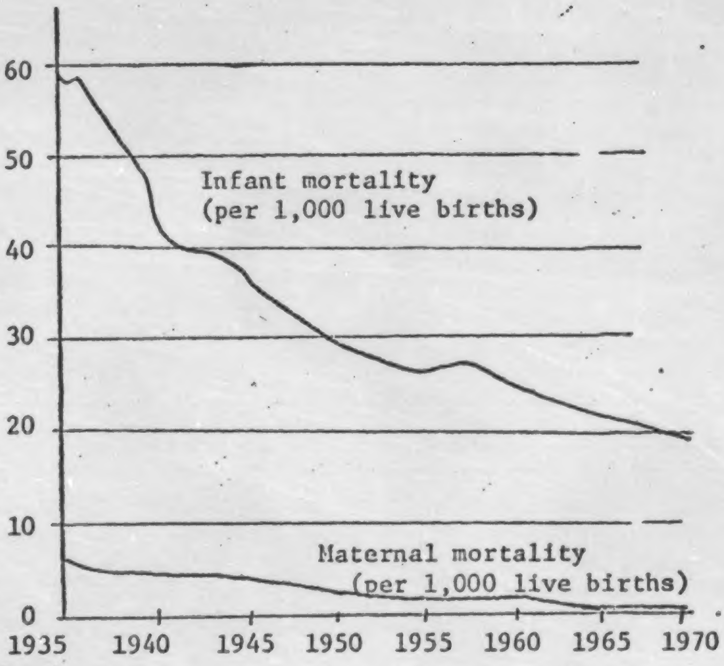


Fig. 3-5 Infant and maternal mortality rate in U.S.

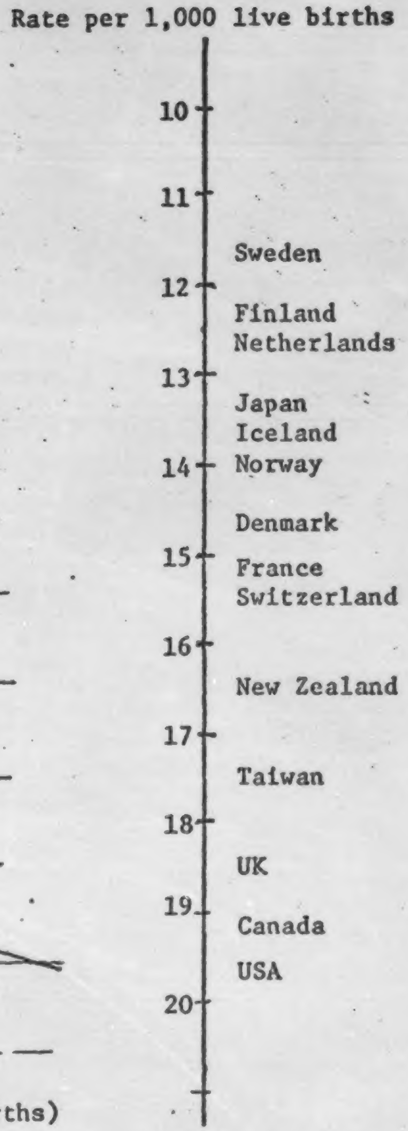


Fig. 3-6 Ranking of countries in infant mortality rate. (from "United Nations Statistical Yearbook, 1971")

As in all major, social problems, there is no simple explanation. A variety of factors have combined to lead to the overall results. In this section, we will look at a few of the explanations which have been presented in recent years by politicians, health officials, and observers of the health system.

(a) "The total number of doctors is too small."

The United States has about 325,000 practicing physicians (one for every 615 people). Sweden (our standard for comparison) has 10,000 physicians for a population of eight million, or one M.D. for every 800 people. Thus, on the basis of the Swedish standard, the U. S. has a surplus of 75,000 physicians. Actually Sweden recognizes it has a shortage of doctors, and the current educational program is designed to bring the ratio to 1 doctor per 500 people by 1980 -- or somewhat better than the current U.S. situation.

Another measure of the adequacy of the supply of doctors is the average number of visits to a physician per year by the citizens. In Sweden, this is less than three visits per year; in the U.S. about five visits per person per year.[1]

These data would seem to indicate that there is no critical shortage of doctors in this country, although an apparent shortage may be created by our tendency to use physicians too much. These figures worry some medical educators who foresee the possibility of an actual over-supply of physicians in another decade as a consequence of the current expansion of medical schools.

(b) "The geographical allocation of physician resources is inadequate." While we may have enough physicians to care for our population, they are

---

1. What constitutes a "visit" has to be defined here, particularly when the patient is hospitalized, but the data for the two countries seem to be comparable measures. The number seems high, but it includes numerous visits by chronically ill patients.

improperly located, with certain segments of the population very well cared for and other segments badly neglected.

In the United Kingdom, a physician is forbidden to set up practice in a geographical area which is already adequately served, and special incentives are given him to locate in rural areas where there are acute shortages of doctors. In the Soviet system, both compulsion and financial incentives are used in an attempt to ensure a reasonably equitable geographical distribution of doctors. In the United States, we historically have assumed that normal economic laws will work; it is only in the last few years that the federal government has developed plans to assist students through medical school, for example, if they agree to practice in areas short of physicians.

Throughout the United States, there are 134 counties (total population of about 1/2 million) which have no physicians. There are many more counties where the ratio of population to doctors is far above the 615 national average, and often where the only practicing physicians are approaching retirement.

(c) "The professional allocation of physician resources is inappropriate."

The argument is that there are too few general practitioners, too many surgeons (as one example of a medical specialty which is over-populated). Medical students tend to drift steadily toward research or specialization as they move through the years of medical school. They are undoubtedly attracted by the excitement of work at the frontiers of medical knowledge, as well as by higher salaries and prestige.

The situation in surgery seems particularly serious. The most common operation in the United States is still the tonsillectomy, in spite of the fact that studies indicate the great majority of such operations are probably unnecessary. (In West Germany, the appendectomy is the most common for some unexplained reason; possibly this is an unstable system in which students are taught to perform operations their teachers favor).

The National Bureau of Economic Research made a small study to investigate the work load of surgeons. In one city, they studied the yearlong records of the 17 surgeons. In order to find a common measure of work, they classified each operation in terms of HE's -- hernia equivalents, based on comparison with a simple, unilateral hernia operation). Various types of surgery ranged from 0.2 HE to 12 HE's, with the evaluation including pre- and post-operative care. The past year's work for each surgeon was then evaluated, and each doctor's work load was measured in HE's per week.

In-depth interviews with surgeons established that a normal, full work load would be 10 HE's per week. In other words, surgeons themselves estimated that a reasonably busy colleague could handle this much work while providing excellent patient care and attending to all ancillary aspects of his profession.

The startling result of this study of one city was that only one of the 17 surgeons exceeded the normal work load; he averaged 13.3 HE's per week. One other surgeon worked just under the normal rate. The others were far below, so far that the average work load was only 3 HE's per week. Surgeons were working at only 30% of normal capacity.

Obviously a study of one city is not necessarily indicative for the nation, but the report of this research reflects a sentiment often expressed in discussions of the state of health care in this country. Since the federal government supports medical education so heavily, there is currently consideration of the possibility of using the force of that financial support to urge medical schools to discourage those specializations in which there seems to be a surplus of professionals.

(d) "The health care of the non-white population is inadequate." The mal-allocations of physicians seem particularly to limit health care for the non-white of the population, in both rural and central-city areas. In one section of the Bronx in New York City, for example, 25 years ago there were 25,000 residents (mostly immigrant families) with only 10 practicing doctors (a poor ratio of 2,500 people/doctor). Today the population is primarily Black and Puerto Rican, numbers 50,000 and is served by two doctors -- a ratio of 25,000/1. Nationally, estimates have been made that 30,000,000 people have access to medical care only in emergency or acute situations.

Part of this difficulty arises from the fact that only about 2% of the physicians in this country are Black, compared to 11.4% of the population. Medicine is an unusually successful profession in attracting minority students, with a 10% figure now realized. In spite of this national effort, the total percentage of minority professionals will grow only very slowly in the future just because of the low starting point.

A more poignant indicator of the problem of health care for minority groups is provided by the life expectancy data of Table 1 and the data on infant mortality. In the latter case, the average for the non-white population is nearly twice that for the total.[1] In terms of the mortality rate of women in childbirth, the U.S. ranks below 100 other nations, and the non-white mother has a rate four times greater than the white.

While the total explanation of the high infant mortality among non-whites is not clear, it does seem apparent that a major element of the problem is that sizeable portions of this population are effectively separated from the health care system. Just three years ago in New York City, as an example, about 40% of the women delivering babies had not seen a doctor during pregnancy. Presumably most of these women knew they were expecting; they did not enter the health care system until it was a necessity.

Why are these people "separated from the health care system?" Again there seem to be several social problems. Many people can not afford the needed medical care and governmental programs are inadequate to cover costs; they don't know how to obtain free health care; they may not know how to use the city transportation system to reach medical care; they may have a deep fear of the health system; and, finally, they may not realize the need for medical care.

---

[1] This situation is a major reason why infant mortality is worse in the District of Columbia than in any of the 50 states, even though D.C. has the most physicians per population.

In this section, we have looked at four explanations for the apparently poor performance of the U.S. health delivery system. Three of these seem to be valid:

(1) The poor geographical allocation or distribution of the supply of doctors.

(2) The inappropriate professional allocation of physicians.

(3) The separation of major segments of the population from the health care system.

In addition to these obvious causes, there are other broad reasons. For example, the population of the United States is a mixture of races; as a result, childbirth (for example) often presents severe medical problems because of the physical differences between mother and father.

5. Male Health

One startling aspect of the life-expectancy data shown previously is the relatively poor situation for males in the United States. Figure 3-4 indicates that the present life expectancy of infant girls is seven years greater than for boys. In contrast, in many of the comparable nations, the female has only a small advantage over the male.

The disadvantaged status of males is a relatively recent phenomenon in this country as shown in Fig. 3-7. In 1901, the male/female ratio varied between 1.0 and 1.2 with age, and there was relatively little difference between the two life expectancies.

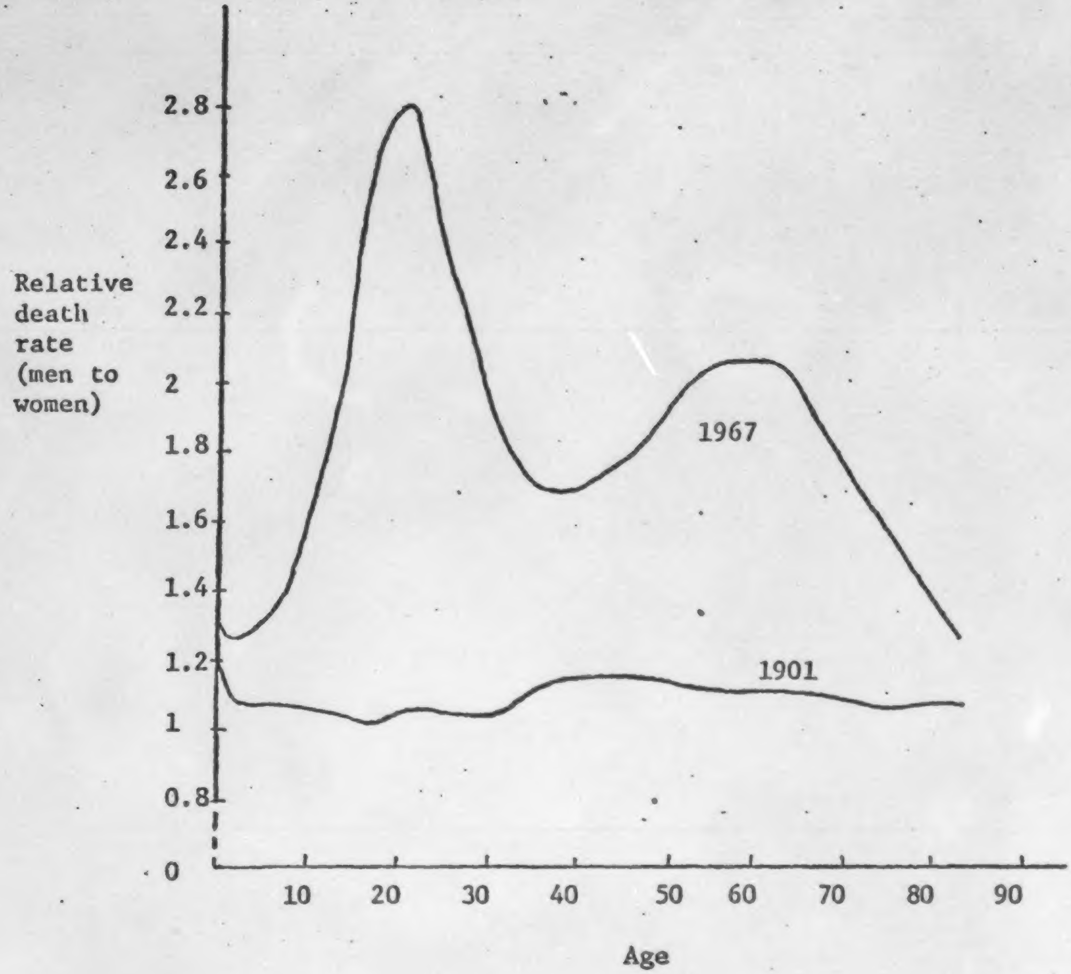


Fig. 3-7 Fraction of men dying compared to women at various ages in 1901 and 1967

Comparison with a country as similar to the U.S. as the United Kingdom demonstrates that the high male, dying rate is an American phenomenon. Between the ages of 20 and 55, the male in the United States has a much larger probability of dying than in the United Kingdom, although after 60 both males and females are much less likely to die than in the UK (one deduces that he should move around the world at each age to find the "best" place to live).

There are many reasons for this poor situation of males. Accidents, suicides, and homicides account for a large fraction of the extra male deaths. This fact is portrayed by a comparison of Figs. 3-7 and 3-8; in the latter case, the rates are shown after the deaths due to accidents, suicides, and homicides have been subtracted. The obvious change is the disappearance of the large peak at age 20 -- a peak therefore primarily the result of these external causes of death. Actually, the deaths from external causes should be increased by at least 10%, since some deaths are listed officially under other causes even though accidents may be the origin of the problem. In this case, the male peak about age 20 disappears all together.

Cigarette smoking is approximately twice as common among males as females (about half the adult males smoke). While the relation between smoking and earlier death is not clear, the best estimate of the effects of smoking is shown by the dashed line of Fig. 3-8. This curve indicates that the ratio, male/female dying rates, still peaks around age 60 after accidents, homicides, suicides, and smoking are removed.

What other factors might account for the relatively poor life

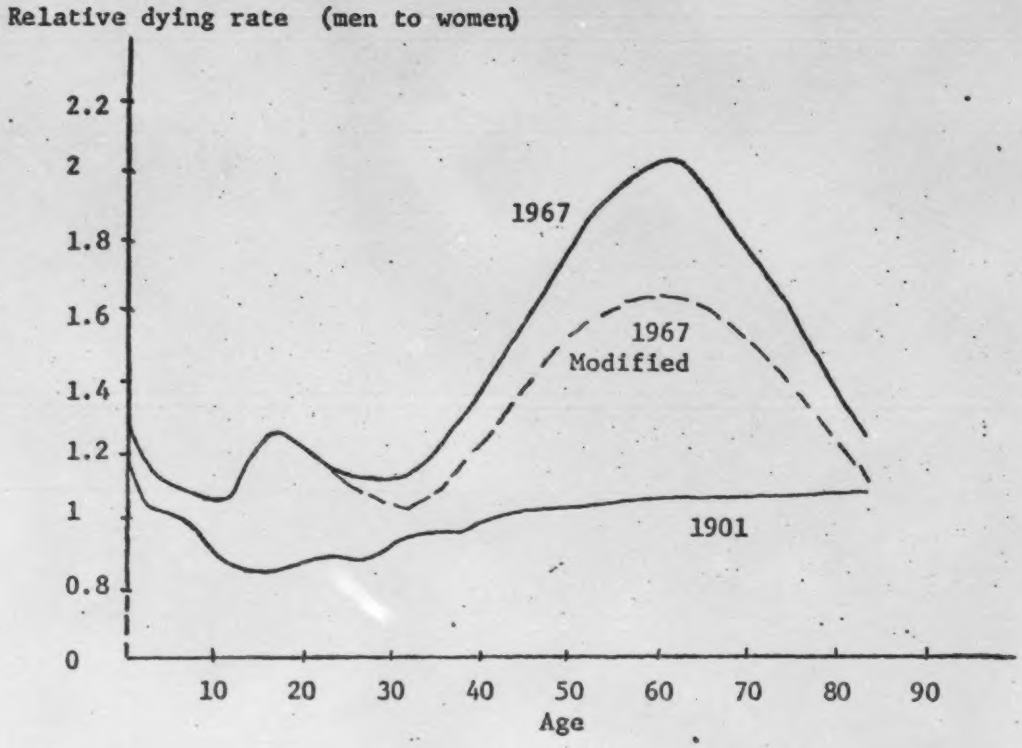


Fig. 3-8 Fraction of men dying compared to women after accidents, homicides, and suicides removed. 1967 modified curve also removes effects of cigarette smoking

expectancy of U.S. males? While the answer is not known, the following explanations have been proposed.

(1) Males are exposed to more unfavorable environments, particularly during working hours. Serious air pollution and high noise levels, for example, seem to be related to respiratory and cardiovascular problems, respectively.

(2) The stresses of working life may be harmful to health.

(3) Alcohol-related problems are discussed in detail in Chapter 4. Of the 60,000 alcohol-caused deaths in 1967, 42,300 were male.

(4) Coffee has been linked to cancer of the bladder and, in recent studies, to "heart" or cardiovascular deaths.

(5) Perhaps 10,000 deaths per year are directly related to illicit

drugs; many of these show up in the category of accidents, suicides, and homicides.

Thus, a variety of factors particularly characteristic of the American way of life seem to combine to cause the poor life expectancy of U.S. males. As Fig. 3-4 shows, female life expectancy in the U.S. is close to the best in the world, while data for males places the U.S. far down in an ordering of nations. Fig. 3-7 displays dramatically the relatively poor status of males; Fig. 3-8 is an attempt to determine a few of the causes for the deterioration since the beginning of this century.

6. Central-City Health Problems

The United States frequently neglects or at best under-emphasizes the health problems particularly important in the central city. We hear the complaint that drug addiction has only become a deep national concern since the "epidemic" has spread to the suburbs.

In the early 1970's, lead poisoning[1] among central-city children has been well publicized. Lead poisoning is particularly dangerous because of the cumulative effect over the years; once deposited in the body, the lead is not normally removed. Excessive lead poisoning affects the liver, kidney, brain, and central nervous system. Children exhibit mental retardation and convulsions, coma, paralysis, blindness, and death.

1971 estimates are that 400,000 children are seriously poisoned each year, 30,000 in New York City alone. Of these, 200 die, 800 become blind, and 3200 suffer permanent brain damage. Six percent of Chicago children in low-income areas have been found to suffer from lead poisoning; St. Louis has 3600 children being treated. Health professionals have deduced that lead poisoning, even at a low level, may be a major factor in the lack of achievement of many central-city children.

---

[1] Lead poisoning was the principal theme of the June, 1971 issue of Environment magazine (Vol. 13, No. 5).

What is the source of the lead? For children there are two major sources, paint and gasoline, with only the latter primarily affecting adults.[1]

Paint

Most of the paint used before 1950 was lead-based. In the apartment houses and buildings where this paint is now peeling off, it serves as an attractive"candy" with a taste like lemon drops. While eating lead is not an effective way to increase body levels (since almost all is excreted from the body), when this input is combined with that taken in through the lungs, high levels can be achieved in even very young children. This seems like an easy problem to solve -- all we need to do is repaint all building interiors (and accessible exteriors) which have not been painted since 1950 (or better, which show lead-based paint). On December 31, 1970, Congress passed the Ryan-Kennedy Bill authorizing thirty million dollars for this purpose; the bill was signed on January 14, 1971. In the ensuing months, very little has happened. Why?

---

1. It has recently been discovered that the paint on the common pencil often contains enough lead to present a serious health hazard for the child who likes to chew the pencil. The Pencil Maker Association, representing the companies making most of the U.S. pencils, has started a program to test each organization's paint every year in an attempt to remove the problem before any serious effects are found.

To understand the inaction, we must recognize that there are two types of bills Congress must pass to ensure a new federal program. First comes an authorization bill, permitting the government to spend (in this case, 30 million dollars) for a particular program. Even after this bill is signed and becomes law, no money can be spent. Next must come an appropriation bill, actually giving the money to the agency or department which is to carry out the program. Only when this bill is also signed can the government actually start the program. In the lead-paint program, only a couple million dollars were actually appropriated.

This two-bill system can serve as somewhat of a means to delude the public. A member of Congress can vote for a very large authorization, then vote against any appropriation. He is then on record as favoring the program strongly, even though he did vote against spending any money (presumably he can argue that he is then fiscally responsible and anxious to keep taxes low).

Furthermore, before a program actually starts, the Executive Office (the President and his powerful Office of Management and Budget, OMB) must agree to freeing the money. In rare cases, Congress passes an appropriation for a Department or agency with the stipulation in the law that a certain amount of money must be spent on a particular program, but usually the President and his staff are given freedom to spend within the limit of the appropriation.

Thus, in the lead-paint program, Congress authorized thirty million, then appropriated far less. The Executive then decided to postpone entirely the start of the program until the next fiscal year. Through this combination of Executive-Legislative interplay, there was no federal action for two years.[1]

Even if the federal program had been fully carried out, it was discovered in the summer of 1971 that the results might not have been total success. The New York City Bureau of Lead Poisoning Control, motivated by the lead-poisoning death of a leopard at Staten Island Zoo, decided to test commercially available paint (after a three-year pause during which no testing had been done, because in 1968 all paints tested were found to have less than the legal limit of 1% lead). It was found that 10% of the paints tested had excessive lead content, with some ranging as high as 10.8% (similar to the level found in the pre-1950 paints). This situation emphasizes the necessity of continual monitoring of products if federal programs are to be effective.

Gasoline

The major source of lead poisoning for most of the urban population comes from the air. One third of the typical, urban adult's body lead comes from breathing city air, one third from eating city food exposed to city air, and the final third from the food itself. The average American now has 1/4 of the lead level corresponding to poisoning, and the average city dweller possibly twice as much. By far the largest contributor to the lead in city air is the tetraethyl lead used in auto gasoline since 1923. Autos discharge 500 million pounds of lead into the air each year in the United States. Unfortunately, almost half of the lead entering the lungs is retained in the body.

---

[1] The process of federal implementation is described in more detail in the appendix of this chapter.

As a result of these facts, emphasis has been placed by environmentalists on encouraging positively the use of unleaded gasoline in cars. New cars are being built by all manufacturers to operate properly on unleaded gas, but the proposal to tax leaded gasoline to raise the price to that of unleaded gas has not been passed by the Congress. While publicity emphasized the dangers of this pollution of the air environment, it apparently was the opinion of the Congress that the public was not yet prepared to pay the extra cost associated with the removal of the problem.

Other urban problems

Lead poisoning is simply one example of those health problems which are particularly important in the cities, even though they are actually national problems. Other examples are sickle cell anemia (affecting primarily the Black population), drug addiction (which only a few years ago was primarily an urban problem, but now represents a nationwide epidemic), alcoholism, and venereal disease.

In the last category, Fig. 3-9 shows the history of reported cases of gonorrhea and syphilis in New York City. (Again, New York City is used as an appropriate example; the present urban growth rate will lead to 12 cities the size of New York today by the year 2000, so New York provides an indication of the problems of the future). The graph shows the sharp decrease in the decade after World War II resulting from the availability of penicillin and the strong publicity campaign to educate the public.

By 1955 venereal disease seemed to be pretty well under control. As so often happens, public interest then diminished and public health measures were relaxed. Since then, there has been a nearly steady rise

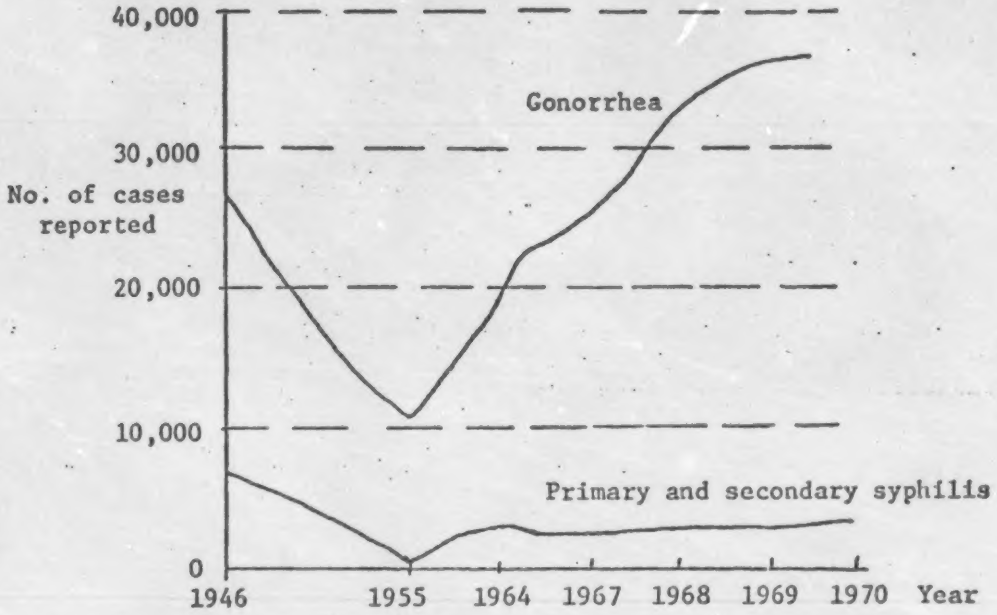


Fig. 3-9 Increase in gonorrhea and syphilis in New York City

until, in 1970, there were 37,000 reported cases of gonorrhea and 4,000 of syphilis in New York City. Estimates are that at least three times this many people actually have the diseases, even though detection is simple and, if the trouble is caught early, cure is also a simple matter.

7. Nutrition

In any study of the health of a nation or a large group of people, we always must wonder to what extent the nutritional standards and customs are a factor.[1] Certainly a comparison of the U.S. and Japan, for example, must take into account the differences in general eating habits in the two countries.

---

[1] Jean Mayer, Toward a National Nutrition Policy, Science, April 21, 1972, Vol. 176, No. 4032, pp. 237-241. This is an excellent picture of the current status of the U.S. nutrition problem. Professor Mayer, of Harvard University School of Public Health, has been a leader in bringing this problem to the attention of both government and the public.

The evaluation of nutritional effects on health is a difficult field because of the lack of scientific knowledge, resulting in part from the difficulty of performing significant experiments on people. For example, when worldwide starvation is discussed, we often hear that the real problem in India is the lack of protein in the diet. A basic question is: what is the daily protein requirement of an average person (protein provides the amino acids needed for growth and for maintaining body tissues)? Even such a fundamental question has not been answered scientifically.

The first widely accepted standard for protein needs was published about a hundred years ago by Prof. Lyon Playfair of the University of Edinburgh. Merely by observing the diet given to welfare patients at his hospital, he decided that two ounces per day were appropriate.

This was the accepted standard until about 1900, when a German nutritionist examined the diet of the average working man in his country and found it contained four ounces. Since the German workers seemed healthy, he concluded this was the desirable level. Later, a platoon of U.S. soldiers was given a diet low in protein (around one ounce); they seemed to be healthy, so the required level was lowered.

Thus, even today we do not know how much protein is required to keep a man healthy. The difficulty of nutritional requirements has been emphasized in the space program, where it will be important to know the needs during periods as long as six months. The problems are compounded not only by the difficulty of doing controlled experiments, but also by the variation in needs from individual to individual. This variation was demonstrated in the British program to eliminate rickets by convincing mothers to give their children vitamin D. Once the program was successful, it was found that an alarming number of children were experiencing serious kidney damage because of too much vitamin D. Even though some kidney problems were the result of a mother's over-zealous use of the vitamin supplement ("If a little bit is good, more must be even better"), studies did show that the desirable dosage had to be determined for the individual child.

In recent years, experiments on animals have suggested that nutrition during the pre-natal period and the first two years of life determines the ultimate brain development (specifically the number of brain cells) -- as well as other physical features, such as the tendency to obesity, the cholesterol level the body considers normal, and so on.

The significance of such scientific understanding is not clear. One physician friend of the author has estimated that as many as 1/4 or 1/3 of the central-city, expectant women he examines are mal-nourished to the point of significant effects on the child. Are significant numbers of children limited physiologically by the age of two in the mental and physical levels they can hope to reach in later life?

There are three interesting side issues if we accept the hypothesis that inadequate nutrition (whatever that may mean) up to age two results in a significant limitation of brain development. First, does this mean that a certain fraction of the population can not be expected to achieve normally in school? If so, should the educational system recognize this fact and design special programs for this group?

Second, the arguments over breast-versus-bottle feeding of infants have often emphasized the psychological effects on the baby and the volume adequacy of the mother's milk supply. Perhaps a much more important factor is the nutritional consideration: how can the baby receive the diet which is best nutritionally? Obviously the question can not be answered until scientists can determine in detail what actually constitutes a diet of high quality.

Third, the pre-natal nutritional process becomes very important. Not too long ago, an obstetrician urged his pregnant patient to make great efforts to restrict severely her weight gain during pregnancy -- often to as little as the weight of the child. Perhaps more sensible guidance will be based upon an understanding of the way the fetus is nourished as it depends on the mother's diet.

Questions such as these point out dramatically the far-reaching implications of new scientific knowledge in an area such as nutrition.

Even if we focus on adults, the United States, in spite of its wealth, has a significant fraction of the population for which there seem to be serious nutritional problems. The first major study of national nutrition was made in the 1930's, when it was found that 1/3 of the people were malnourished, a large fraction partly because of the Depression. This survey led directly to such steps as the school lunch program and the enrichment of white bread with B vitamins.

Subsequent studies in 1942, 1948, and 1955 showed an improving situation because of both economic changes and more effective educational programs. The 1965-6 survey, however, brought out certain adverse changes primarily resulting from the increased consumption of saturated fats, salt, and sugar, and the development of food fads and a wide variety of books urging odd diets.

By the time of the 1969 White House Conference on Food, Nutrition, and Health, the wealth of the country had increased to the point where only 16% of our spendable income was devoted to food (in contrast to 30% in western Europe and 50% in the USSR). Yet about 25,000,000 Americans had an income too low to permit a reasonable diet. Furthermore, many people consume in great quantities those foods which have a very low nutritional value, such as candy, soft drinks, and crisp-fried snacks.

National attempts to improve the nutritional picture have concentrated on the

(1) School lunch program (now covering 8 million poor children)

(2) Food stamp program (over 10 million recipients)

(3) Family food assistance, involving food distribution (over 3 million people)

In addition, the White House Conference urged expanded programs of public education, improved labelling and grading, intensified research particularly on the effects of additives, and federal prods to increase food enrichment and fortification.

In the general problem of nutrition, as in all of the areas discussed in this book, the ultimate question is one of priorities. Within a limited total budget, most of which is specified from the outset for on-going programs, the legislative and executive branches of government must select those problems which seem most urgent and in which progress appears likely. For example, the Food and Drug Administration, with its responsibility to monitor and control the food processing industry, must compete with other governmental agencies and programs for available funds; once the budget is established, FDA necessarily has to restrict its efforts to the most pressing areas. While no one would argue with the importance of adequate nourishment for every person, there are always the questions of what programs can be sacrificed to achieve success here, and, even if money is made available, how can a successful plan be designed and implemented.

The difficult role of the FDA is illustrated by the recent history of the DES additive.[1] DES (diethylstilbestrol) is a synthetic female sex hormone which in the 1950's was known to cause the growth of female breasts on a man; and in the last year to have led to vaginal cancer in young women who had been exposed to it as fetuses. DES is a known carcinogen, which has been used for some years to hasten the fattening of cattle.

Because of the difficulty of measuring DES residues in beef and uncertainty as to whether there is a safe level below which no harmful effects result, the FDA has allowed the use as a cattle feed additive. For many years, the DES was forbidden during the two days before slaughter; this withdrawal period was then lengthened to seven days. At the beginning of 1973, DES was finally banned altogether in the feed, but still permitted as a pellet placed in the animal's ear.

The conflict arises because the additive represents a saving of about $90 million annually for cattle-raisers. The argument is based scientifically on the question of whether there is a threshold level below which DES does no harm to human beings. Since this threshold is not known (or may not even exist), the opponents of DES argue that the chemical should be banned totally, even if the result is a rise in the cost of beef to the consumer.

---

[1] Science magazine: July 28, 1972 (335-337); August 11, 1972 (503); August 18, 1972 (588-591)

8. Management of Health Delivery

In the preceding sections, we have looked at the U.S. health picture in an attempt to begin to understand some of the reasons for the relatively poor situation in this country. From this very brief and limited study, it seems apparent that a significant fraction of the population does not have adequate access to the health care system. Furthermore, there is too often no concerted, national program to conquer specific problems. While as a nation we spend large amounts of money on health, there is relatively little centralized, logical planning of the "campaign" against disease, malnutrition, and physical disability.

There are two ways in which this absence of logical planning is apparent. One is in hospital use. If one examines the hospitals in this country, he finds a wide range of crowdedness -- in some locations, patients have to wait for months for admission for routine treatment; in other areas, hospitals are not filled.

In the crowded hospitals, there are many reasons for the overload. In some cases, patients are kept for extremely long periods of time (sometimes because the doctor recognizes that the patient's home environment is apt to harm recovery). Hospital management is often less than ideal.

There is often little coordination and planning among hospitals in a region; it is not unusual to find a half dozen hospitals in one city equipped to do open-heart surgery even though some of them may do less than one operation per week.

In quite a different area, the effect of the lack of national planning appears. All too often, we as a nation seem to be unable to select national priorities and then pursue these to a successful conclusion.

For instance, by 1971 artificial kidney machines have become standard for treatment of patients in whom the kidneys do not function properly. In many cases, regular use of the kidney machine to "cleanse" the blood permits the patient's kidneys to recover, although treatment often lasts for years.

While there are attempts to develop a portable artificial kidney which the patient can use at home, almost all treatment is now done in a hospital where the machine is shared by a number of patients. The cost per patient runs on the order of $20,000 per year, so the treatment is hardly trivial. Unfortunately, the national supply of the machines is severely limited, and estimates have been made that 30,000 people/year die because they are unable to obtain use of a machine. Many of these people are in their productive years.

A national program to purchase and operate machines for this group would cost perhaps $120,000,000 per year -- certainly not an insignificant amount, but still small compared to the total health expenditures. Why has such a program not been adopted by 1973 (although one was approved in the closing days of the Congressional session in 1972)?

Part of the difficulty certainly comes from the way national priorities are established. A national program can be launched by the President (as Kennedy's decision to land men on the moon within a decade), but he must solicit public support to ensure the appropriations by Congress. One way to gain popular and political approval is to assemble a group of health and medical experts and hope they will endorse the program enthusiastically.

Unfortunately, when a group of health authorities is assembled and they are asked what the nation should do if an additional $120,000,000 a year were available for health, there is rarely any agreement. The heart specialist emphasizes that his concern is our greatest killer: the cancer specialist that his area is the major public fear; the preventive medicine expert is concerned with the overall health benefits from an urban program: and the ophthalmologist points out that as a nation we have done very little to make life more productive and pleasant for the two million blind citizens.

The result all too often is a small amount of support for each of many different programs -- with none of them pushed hard enough to yield significant results.

9. Conclusion

In the health field, medical knowledge and technology are now at the stage where major progress can be made if this nation develops a logical plan which encompasses education of the public on both health and political issues, forceful programs aimed at specific problems, and intelligent management of major elements of the health delivery system.

Furthermore, in spite of the statistical measures showing the U.S. to be lagging certain other advanced countries, there are other data which are an impressive testimonial to American achievement. For example, in the United Kingdom, in spite of socialized medicine, dental care is deplorable: 30% of the Englishmen over the age of 30 have none of their own teeth. In many aspects of health, the U.S. offers unparalleled care; the only real problems are modifying the system so that a larger proportion of the population has access to this care, and responding to the changes in health services needed by the population.

In this latter category, the situation is particularly critical with respect to one facet of the health problem which has not been mentioned in this chapter -- the senior citizen. The population over age 65 is increasing rapidly, measured as a total or as a percentage of the total (Fig. 3-10). More than 3/4 of these people suffer from chronic disease, and the group places very heavy demands on the system for both physical and mental health services.

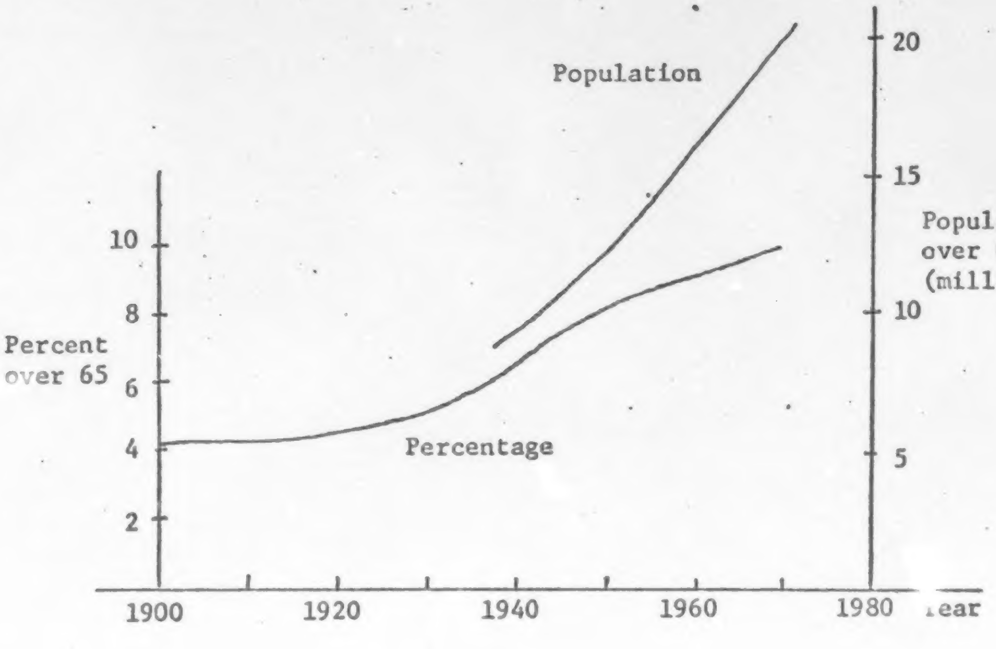


Fig. 3-10 U.S. population over age 65. (In under-developed countries the percentage is 3-4; in 1963 in France it was 12% because of widely varying age distributions resulting from two world wars).

APPENDIX 3-1

FEDERAL IMPLEMENTATION

Even after a national policy has been agreed upon, putting that policy into practice may not be a simple matter. A program is not effective unless it is provided with money which in turn is actually spent. The recent history of the federal government frequently shows publicized and attractive programs in which enough funds just have not been spent to give the effort any reasonable chance of success.

The spending of money on a particular program is normally preceded by a long series of steps spanning a couple of years. A brief description of these steps indicates the various ways in which the best intentions can be thwarted and also some of the techniques by which the voting public can be misled.

In order to simplify the discussion, we will dream up a particular program -- the provision of free hearing aids and tape-recorded daily newspapers for citizens who are hard of hearing. The program is called Sensory Orthotics[1] and Newspapers for the Deaf and is known by the acronym SOND (pronounced "sound"). It is administered by HSMHA (Health Services and Mental Health Administration) of HEW (the Department of Health, Education, and Welfare). Money is divided between furnishing materials to deaf people and research on new hearing aids and other innovations to assist the hard-of-hearing.

---

[1] Orthotics are devices to supplement the human system, while prosthetics replace human parts. An artificial leg is a prosthetic device, a hearing aid or eyeglasses is an orthotic device.

We wish to trace the steps through which dollars are made available and actually spent under SOND. We focus on the year from July 1, 1973 through June 30, 1974. The federal fiscal year starts July 1; the year ending June 30, 1974 is called Fiscal Year 1974, or FY'74.

(1) Spring 1972. The men running SOND prepare their estimate of the dollar needs for FY'74. The program is reviewed and plans are made for two years into the future. The outside advisory committee, composed of experts on education and devices for the deaf, works with SOND personnel and has a series of meetings to advise on desirable programs. The SOND personnel and their bosses in HEW meet with staff members of both OST and OMB in the White House to consider the relative importance of SOND compared to other governmental programs.

These two agencies, OST and OMB (the Office of Science and Technology, and the Office of Management and Budget), are part of the large staff of advisors of the President. The OST, directed by the President's Science Advisor, is the principal source of information on scientific and technological matters. The OST staff consists of about 30 scientists and engineers, each responsible for the total governmental program in one specific area. The staff is advised by PSAC, the President's Science Advisory Committee.

OMB is one of the most influential groups in the executive branch of the government, as we shall see in the following paragraphs. In broad terms, OMB is responsible for the fiscal operation of the government. Following the guidelines laid down by the President and his immediate advisors, OMB puts together the budget each year. After the moneys are appropriated, it controls expenditures by releasing money to the various government agencies. OMB consists of a large staff of professionals, each an expert in one facet of the federal program.

(2) Summer 1972. After the planning meetings of the spring, the FY'74 budget of SOND is put together during the summer. From the discussions, the leaders of SOND have obtained a general idea of the acceptable size for the budget. If SOND is viewed as an important effort by the Executive branch, for instance, a budget increase can be considered -- perhaps by as much as 25%.

(3) Fall 1972. The SOND budget is combined with other parts of HSMHA, and then with the total picture for HEW. Detailed documents are prepared to describe the total HEW program, with SOND merely one small part. The HEW secretary frequently discusses the total departmental picture with the President and White House staff.

(4) January 1973. In his Budget message to the Congress, the President presents the total federal budget along with broad descriptions of the primary goals of the FY'74 program. OMB prepares the detailed budget picture, by agency, for forwarding to the Congress, and also a breakdown of the budget by category. In the last document, the SOND program is part of health services and also of biomedical research; the HSMHA efforts in these directions are combined with those of the Veterans Administration, The National Institutes of Health, The Bureau of Indian Affairs, and so on, in order to give a total picture of the federal effort in health care.

(5) May 1973. Action now switches to the Congress. In both the House and the Senate, there is a committee with responsibility for authorization bills for the health efforts of HEW. The first step through Congress is this authorization (we assume SOND is not a new program requiring new

legislation). Each committee holds hearings on the authorization. Members of HEW staff (including the director of SOND) testify, as well as invited experts from outside the government. Unless the program is controversial, these hearings in both the House and Senate are likely to be reasonably friendly affairs. Usually, lengthy statements are read into the record before only one or two members of the committee (Congressmen and Senators carry a fantastic work load and just are physically unable to attend any but the most important hearings and meetings).

(6) May 1973. An authorization bill comes to the floor of the House. Often this authorization is for appreciably more than the President's budget, as the House committee adds to programs it finds attractive and important. The House then approves the authorization after debate. This is the ideal time for Congressmen to go on voting record in favor of the programs. This is only an authorization action, not an appropriation. No money is actually being spent or even allocated. The House is merely saying that it is authorizing subsequent appropriations up to a certain figure (say $20,000,000) for SOND. The President only asked for 10 million in his budget; Congress has not yet appropriated anything; it is stating that its future appropriation will not exceed 20 million.

Thus, when a Congressman running for reelection claims he voted for the SOND program (while giving a speech at a school for the hard of hearing), one must ask whether he is referring to his vote on the authorization or on the appropriation. Certain Congressmen have voted for authorization, then against any appropriation at all.

From the House, the authorization goes to the Senate where, again after hearings, a bill is passed. If the House and Senate bills differ, a joint conference committee is established to develop a single, compromise measure, which is then passed and sent to the President. If he signs the bill, the authorization process is completed.

(7) Summer 1973. Before the authorization is done, a different House committee starts consideration of the appropriation bill. Now we are dealing with real dollars which Congress is assigning to the Executive branch to be spent. The committee and House must take into account the total federal budget picture and must make difficult choices among alternative programs, each costing money.

Hence, the House appropriation for SOND is likely to be well below the authorization -- perhaps 12 million dollars (compared to the 20 million authorization and the 10 million in the President's budget). As indicated above, at this stage we may even discover that the entire appropriation for SOND is killed, or that the funding is so small that the program has no chance to succeed.

Again, after passage by the House (according to the Constitution, all appropriation bills must originate in the House of Representatives), the bill goes to the Senate where the same committee procedure is followed. The appropriation is not completed until a single bill is approved by both houses and signed by the President. If the total appropriation for HEW is too different from the budget proposal, the President is likely to veto; if Congress does not over-ride, the entire appropriation process has to start again.

The appropriation should be completed by July 1, 1973, at the start of the fiscal year for which money is being provided. In many cases, the appropriation is not made until the fall. If no action is taken by Congress, all programs terminate and staff can not be paid. Consequently, once July 1 passes, Congress usually passes a resolution which allows all such programs to continue spending at the rate of the preceding year until a fixed date in the fall or until the new appropriation is approved.

(8) Fall 1973. We are now into the FY'74 and Congress has included $12 million for SOND within the total HEW appropriation. SOND still can not spend the money, however, until it is released by OMB, acting for the Executive. OMB may decide to release only $2 million to carry the program until November 1, by which time HEW should provide a detailed statement of the long-term goals for SOND.

Regardless of the particular procedures used for obtaining Executive approval of the actual use of the money appropriated, OMB can determine the amount released to SOND for the year. If OMB wishes, this may be only the $10 million originally requested in the President's budget; indeed, if federal expenditures rise alarmingly during the year, OMB may decide to cut SOND back to $8 million.

Thus, the Executive branch (through OMB) has the final determination of the level of spending for SOND. In rare cases, Congress may insert into the appropriation bill the requirement that no less than a specified amount must be spent on a particular program. This form of congressional control over the executive powers is only used for very important programs in which there is a history of sharp differences between the executive and legislative branches.

Final comment.

The significance of a program (such as our SOND example) depends most strongly on the funds actually spent. These in turn are determined by a long sequence of decisions in both the executive and legislative branches -- decisions starting with the attitudes of higher officials in the home Department (here HEW). If a significant expenditure increase is to be obtained for SOND in FY'74, the director and proponents must convince:

Higher officials in HEW

White House staff members, particularly in OMB

Members of the four Congressional committees (two in each chamber) concerned with authorization and appropriation

There exists a long sequence of points at which the total dollars available can be cut; restoration of funds is increasingly difficult as we move down this chain covering two years in time.

CHAPTER 4

EMERGENCIES

Accidents, whether fatal or serious, have a strong impact on the quality of life because they by definition occur unexpectedly and they often affect young and healthy people. The manner in which a society cares for its accident victims is one measure of its technological sophistication, for the quality of this care is determined largely by the resources which the society is willing to devote to helping the unfortunate minority. This broad statement is even more apparent if we include the physically and mentally handicapped with the accident victims.

The magnitude of the accident problem in the United States is emphasized by a comparison of 1900 and 1960:

	1900	1960
Deaths per 100,000 population, all causes	1719	946
Deaths by accident, per 100,000 population	72.3	51.9
% of all deaths due to accidents	4.2	5.5
% of all deaths due to violence	1.4	1.5
% of all deaths in ground transportation	0.8	2.4

While the number of accident-caused deaths has decreased, the change has been much less than the drop in the death rate. In other words, accidents constitute a much more important part of the total health scene today than they did in 1900.

The two boxed rows show the most important statistics. While accidents, as a cause of death, have increased from 4.2% to 5.5%, the particular category of transportation accidents more than accounts for the change. In 1900, transportation involved horses, vehicles, and railroads; the automobile was an insignificant factor. In 1960, transportation (essentially just motor vehicles) accounted for 44% of all accident deaths.[1]

1. Automobile Accidents

More than 50,000 deaths and 450,000 serious injuries each year are caused by automobile accidents. Meanwhile, the number of motor vehicles in use in the United States continues its rapid growth (Fig. 4-1), which

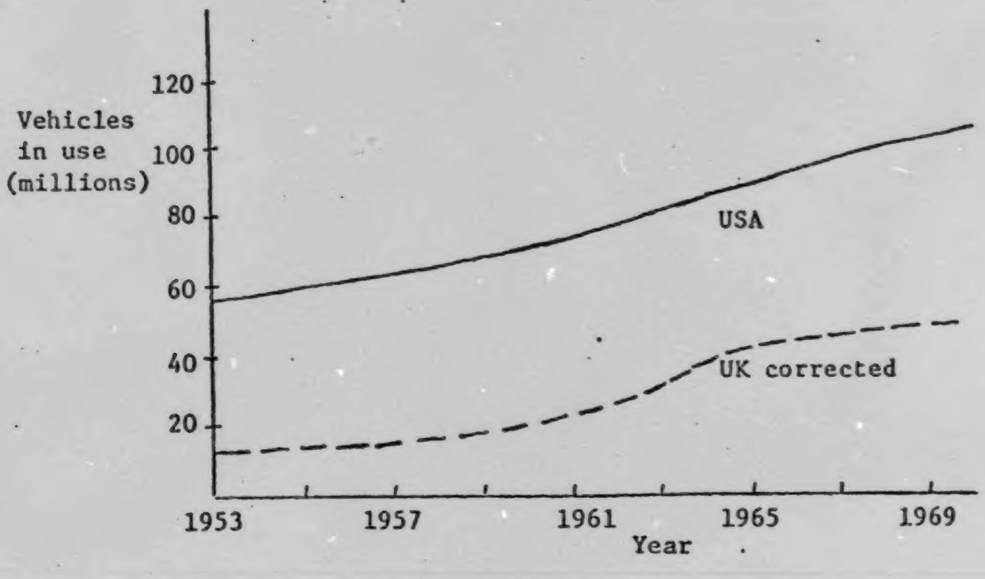


Fig. 4-1 Number of motor vehicles in use in the U.S. and corresponding data for the United Kingdom (adjusted to make total populations equal)

obviously will eventually level off after there is slightly more than one vehicle for every registered driver. Unquestionably a major cause of these accidents is the driver under the influence of alcohol; this is such a serious and distinct problem, it is discussed separately in Sec. 4-3. In this section, we focus on other ways of classifying the causes of accidents.

---

[1] The accident problem is almost as old as the automobile. According to a famous anecdote, in 1892 there were only three cars in the United States. Two of them were in St. Louis, where they were involved in a serious collision.

The fatalities in motor vehicle accidents are often attributed to three separate collisions:

(1) Collision of the vehicle with another vehicle or other object,

(2) Collision of the passenger with the vehicle in which he is riding,

(3) Collision of the injured person with the emergency medical care system.

The estimate has been made that about 1/3 of the deaths can be attributed to each of the three collisions.

Collision-one deaths are those caused by the initial, accident contact -- the "instantaneous" death, for example, from a car at high speed striking an abutment. Collision-two deaths occur when the occupant hits part of his car or is thrown from his car. Collision-three deaths are, broadly, those which could be prevented if the first person at the scene were a qualified physician with adequate emergency equipment.

Collision one

Collision-one deaths can be reduced by optimum design of the highway and traffic-control system and by building extreme reliability and maneuverability into the vehicle. The interstate highway system has been one of the most important factors in reducing the accident rate per vehicle mile over the past 15 years. The divided highways and limited-access roads minimize the opportunities for two-vehicle collisions. Actually, the key concept of a one-way street was introduced in 1760 in the neighborhood of New York's John Street Theater, where "traffic" was hopelessly snarled at the beginning and end of performances. The one-way streets, which have become common in the last two decades within most of the major cities, not only expedite traffic flow, but also reduce dramatically the accident rate.[1]

Furthermore, the new superhighways are being built with vastly more attention to quality engineering, expressed in such terms as low-skid surfaces on treacherous locations, change in surface off the edge of the pavement so that the drowsy motorist is alerted if the car veers, and road curvature and scenery change designed to reduce monotony. Coupled with such construction developments, there is a rapidly increasing use of modern communications equipment to alert the motorist to potential problems ahead. The California Highway Department has wired 42 miles of freeways with sensors embedded in the roadway to measure traffic density and speed. Comparison with computer-stored records typical for the time of day and weather conditions indicates very rapidly any unusual tie-ups. Warning signs can be flashed to approaching motorists. One hope here is to avoid the chain-reaction accidents in which scores of cars pile up after an initial blockage of the road.

---

[1] In the U.S., driving on the right side derives from the use of oxen and horse teams that opened the West. The left front animal was always the leader. When two trains met, the drivers dismounted and hand-led the teams past one another.

In England, driving on the left arose a thousand years ago when reins were held in the left hand, so that the right hand would be free to wield a sword against attacking highwaymen. The result is the driving difficulty today when an Englishman visits the U.S. or an American rents a car in England.

Other developments in highway improvement range over greatly improved lighting and more carefully designed signs. In the latter category, the U.S. is now converting from word signs to the standard picture signs common throughout Europe (where the variety of languages was an early incentive for the change). During the five-year conversion period (starting in 1972), the most important picture signs will be accompanied by the older word signs to allow a gradual education of the motorist.

In 1970, the federal Department of Transportation contracted for the design and development of "experimental safety vehicles" -- cars in which an attempt was made to look into the future and see what novel safety features could be included in a passenger car. The three goals were increased ability to avoid accidents, reduced damage to the vehicle in an accident, and improved protection of the occupant (in other words, reduction of fatalities in both Collisions one and two).

In addition to these major programs, research is currently underway on a variety of specific developments. A periscope is being studied to improve rear vision: the problems are the cost of a large mirror over the roof and, above all, the weakening of the roof by the opening which brings the visual image down to the driver. A vibrator is available on the steering wheel to warn passengers when a horn or siren is sounded nearby -- a device useful not only for the driver who has poor hearing, but also for the motorist with his windows closed and air conditioner and radio on. The problem here is strictly technological: how to build a device which always recognizes a siren but does not respond to similar sounds which may appear from other sources.

In spite of these advances, the safety problem is increasingly difficult because of the tendency of speeds to rise. Many of the parkways only 30 years old and still heavily used were designed for cruising speeds of 35 mph, but now have speed limits of 55. This continuing growth in speed is emphasized by the fact that the first statewide speed limit was 12 mph in the country, 8 mph in the city (Connecticut in 1901).

Collision two

The second collision is the one which is caused by occupant motion after the vehicle collision proper. This second collision includes the motion of the passenger against the interior of the car, the whiplash resulting from different accelerations of his head and torso , and the possible ejection of the occupant from the auto. As indicated earlier, roughly one third of the fatalities can be attributed to this part of the accident (although it naturally is often impossible to assign the cause of death to any one of the "three collisions").

The obvious answer to the problem of second-collision fatalities and serious injuries is the lap belt and shoulder harness combination. A variety of studies has suggested that just over half of both fatalities and injuries could be averted if 100% of the front-seat occupants used the three-point harnesses which are now available in all new cars. Use of just the lap (or seat) belt would reduce the deaths and injuries by more

Table 4-1

Sample of Auto Safety Regulations

- January 1968
  - Padded dashboard
  - Energy-absorbing steering column
  - Recessed knobs on dashboard controls
  - Lap belts
- January 1969
  - Headrests
- January 1970
  - Steering wheel lock
- January 1972
  - Side marker lights
  - Hazard warning lights
  - Shoulder harness for passenger beside driver
  - Warning light and buzzer for front-seat occupants' lap belts
- September 1973
  - Energy-absorbing bumper front and rear
- April 1973
  - Illegal to modify (or disconnect) safety equipment
- September 1973
  - Ignition interlock tied to lap belts
  - Stricter bumper standards and uniform bumper height
  - Strength of side door specified
  - Roof strength at least 1½ weight of car
  - Fire-retardant interior fabrics
  - Tail light improvement
  - Illuminated knobs for wipers, defrosters
- September 1975
  - Passive restraints
- 1977
  - More stringent roof standard

than a third.[1]

In spite of the wide publicity given to the advantages of seat belts and harnesses, usage in this country has increased very slowly. For cars in which the equipment is available but there is no warning signal if the seat belts are not fastened on all occupied front seats, belt usage has apparently stayed around 20-25%. When the warning systems and ignition interlocks are included, lap belt usage seems to rise to about 70%, although shoulder harnesses are still used by less than 10%.

A natural question arises here: If so many deaths and serious injuries could be prevented by the use of lap belts and shoulder harnesses, why not require such equipment by law? In other words, why not fine occupants in the front seat of a moving vehicle who are not adequately restrained? Or decree that insurance is invalid if an occupant is unbelted?

Actually, the fine ($20) has been tried the last two years in Victoria, Australia, with the satisfying result that 3/4 of the riders use lap belts. A few attempts to introduce similar state legislation in this country have been unsuccessful, apparently because of legislators' opinion that the public support does not exist. The idea that insurance coverage should depend on restraint usage is unattractive because of the difficulty of determining, after an accident, whether the occupants were belted or not.

---

[1] The best data here are from a study by Volvo, the Swedish manufacturer. In order to obtain reliable data, the company gave free with each car purchase a collision insurance policy. To collect, the owner had to report the accident in detail. Almost 30,000 accidents were described, involving 26 deaths (25 unbelted) and 2500 injuries. Several studies have shown no fatalities for belted occupants in accidents with vehicle speeds up to 60mph.

The low usage of lap belts and especially harnesses is partly due to the discomfort associated with the devices. Because of the tremendous variability of sizes and shapes of human beings, it is extremely difficult to design a single harness which is comfortable for almost everyone. Equipment adjusted to the individual is expensive.

The protection of infants and small children is a particularly important problem because these passengers are likely to be thrown wildly by a collision. Most of the 1000 children under five killed annually and the 6000 seriously injured could easily be protected by suitable restraints. In 1972, Consumer Union of the United States published[1] its analysis of the efficacy of 17 commercially available child seats and infant carriers (all of which met the government standards which specify withstanding a static pull of 1000 pounds). In actual collisions using, for example, a 30 mph sled hitting a fixed barrier (equivalent to a car at 60 mph hitting a parked car), only the General Motors Infant Carrier passed tests involving front, rear, and side collisions. Two others (Ford ToT Guard and Sears Safety Harness) were rated "acceptable," and an additional two "acceptable but poor." All others were deemed "not acceptable."

Both the lap belt and shoulder harness fall into the category of active restraints. By "active" we mean that protection depends on an action by the occupant: he must fasten the equipment. The most attractive restraint would be one which is passive: that is, one which works even if the rider does nothing himself. In the next section, we discuss the recent development of the air bag, the most promising passive restraint, as well as other auto safety features to decrease the accidents and fatalities in the second collision.

---

[1] August, 1972 issue of Consumer Reports.

Collision three

The third collision refers to the first aid and emergency medical care given to the accident victim. Attributing one third of the fatalities to Collision three is another way of saying that one third of the deaths could be avoided if a qualified doctor and equipment were with the victim shortly after the accident. Clearly, we are comparing the actual situation with an ideal which can never be approached, so that the actual 18,000 deaths (one third of the total) could not be reduced to zero even if the emergency medical system were dramatically improved.

But, as pointed out in Section 4-4, the emergency care system in this country is far from ideal; unquestionably a significant number of lives are lost simply because the emergency care is inept.

2. Auto-Safety Controversy

Ralph Nader's crusade for automobile safety has resulted in a series of U.S. legislative acts. A few safety features were required in new cars in 1968: interior padding, head restraints, collapsible steering column and lap belts. Bumper standards are imposed in 1972 and 1973, flammability standards in 1972.

The most controversial regulations relate to passive restraint systems -- that is, restraining devices which do not require that the rider take any action such as buckling a seat belt. Under the latest law, this safety feature is required by August 15, 1975 (although a one-year delay is currently being requested by the manufacturers).

The auto companies have argued strongly that the air bag is the only device which can meet the 1975 requirements and that the air bag is too expensive and as yet unproven. Furthermore, to provide protection in side collisions, the structure of the car must be changed. In the face of these difficulties and with the air bag not fully developed or tested, the auto manufacturers urge a postponement of the 1975 date. Claims are made that the device will be exorbitantly expensive for the public, that protection will be uncertain (since possibly seat belt usage will drop even lower), and that the system reliability will be unproven, particularly after the automobile has been in use for some time.

The consumer advocates reply that these are familiar attitudes of an auto industry which has long given little real attention to safety, that the costs are small compared to the frequent auto style changes which serve no purpose except to induce people to buy new cars, and that only government action will ever force the industry to strike an appropriate balance between safety and appearance.

In an attempt to evaluate these vehement arguments, the Office of Science and Technology (OST) in the Executive branch established a special committee in 1971 to consider what federal safety regulations are reasonable. What devices can be required by 1975 without excessive increases in auto cost or a poorer performance? What are the alternatives available to the public in this trade-off between safety and, on the other side, cost, appearance, and performance?

The committee reported on March 19, 1972 in a book entitled

"Cumulated Regulatory Effects on the Cost of Automotive Transportation." The document was labelled the RECAT report from the underlined letters. This report is an intensive study of two facets of regulation of automobiles: air pollution and safety. In the latter area, the report attempts to evaluate the developments of the preceding five years in the field of auto safety: the highway improvements, the changes in car design, and the various laws enacted.

The report recognizes that, while new safety features are certainly desirable, there has to be a compromise between safety and cost. If car prices go too high because of mandated safety features, sales will drop and there will be a serious impact on the U.S. economic and employment picture. Furthermore, safety devices should not be required before they are proven, or at least until there is a reasonable chance of success.

A day after the report was made public, Ralph Nader issued a statement, the first paragraph of which reads: "On behalf of the White House, the Office of Science and Technology has prepared and released the RECAT report as another effort to intimidate the federal regulatory agencies responsible for regulating motor vehicle air pollution and safety. Never in the history of the Office of Science and Technology has it been so manipulated or, in the current parlance of the White House scene, so Flaniganized. This report was the inspiration of Peter Flanigan who pursued it all the way to its predetermined destination. It is a mockery of scientific integrity and competence and a penny-ante caricature of think tank studies."

What is the reason for this sharp controversy between Nader's group and the White House staff? Certainly both are interested in improved auto safety.

In this brief section, we can not discuss any significant fraction of this Report of several hundred pages. Rather, we will look in detail at merely one argument, that of the air bag, and focus on the differences of opinion with regard to the speed with which the air bag should be required in all new cars.

Air bags

Under the current regulations of the Department of Transportation, all new cars sold after September 1975 must provide passive protection for all occupants in front, side, and roll-over accidents at speeds of 30, 20, and 30 mph, respectively. We should emphasize here that the system required is passive: i.e., no action is required of the occupant. Hence, lap belts and shoulder harnesses are not sufficient.

According to the auto manufacturers, the only way to meet this standard is with the air bag. This is a porous cloth bag normally folded in front of the occupant. The bag is connected to a supply of nitrogen or another gas at 2000 pounds/square inch pressure. Sensors determine when a serious impact occurs (by measuring the large acceleration immediately following a collision). When the sensor detects a bad impact, gas is released into the bags. The bag must inflate in 1/20 second to protect the rider. Within seconds thereafter, the bag deflates. In other words, the inflated bag in front of the rider gives him a "cushion" into which he can move without hitting the dashboard, windshield, or steering wheel.

The air bag is still an exceedingly controversial issue. The government originally planned to require installation in 1974 and 1975 model cars, then backed off to postpone required introduction until the 1976 models (that is, from September, 1975 on). The auto makers now urge an additional year postponement.

The arguments for and against the air bag are easy to list, although it is difficult, if not impossible to decide which are really valid:

For

(1) Passive restraints (that is, the air bag) are essential since Americans will not use lap belts and shoulder harnesses. The air bag plus the harness is by far the best approach to safety.

(2) Current laws require a buzzer-light warning system and ignition interlock tied to the lap belts for all front-seat occupants. The lap-belt warning can be cheated trivially; one needs only fasten the belt and then sit on it. In addition, the entire system is easily deactivated.

(3) The estimated cost ($300 per car) is small compared to the cost of style changes which are useless and primarily serve as customer inducements to purchase a new car.

Against

(1) The cost is likely to be as much as $500 per car, which will have to be passed on to the purchaser whether he wants the air bag or not. Car sales will plummet (as people postpone a trade-in for several years), and employment in all related industries will suffer.

(2) The lap belt-harness system with ignition interlock has not really been given a trial yet. Since the average life of a car is ten years, any innovation on new cars must be followed for nearly ten years before half the cars on the road are equipped.

(3) When air bags are installed, the use of belts and harnesses

will drop.

(4) The bag may explode accidentally while the driver is cruising at high speed.

(5) The explosion of the bag causes a noise which may be damaging to the hearing of car occupants.

(6) Reliability is undetermined, particularly after years of car use; mechanics are not trained in maintenance.

(7) The sudden blowing-up of the bag results in high pressure within a closed car which may push the doors open, break a window, or even force gas out of the tank.

Obviously, the validity of such claims needs to be determined before a decision is made. One of the problems in scientific testing is that the air bag is a very new device; only a few engineers (normally employed by the auto makers) are familiar with the current status of the development.

In an attempt to gain some experience with air bags, Allstate Insurance Company in 1972 purchased 200 Mercurys, each equipped with an air bag in front of the passenger in the front seat. These cars were used by executives of the company, except for two which were allocated to collision tests with both dummies and a human passenger. These tests were made in the course of producing television commercials to introduce the air bag to the American public. Unfortunately, all three TV networks refused to broadcast the commercials on the ground that they dealt with a controversial topic. Although a few local stations did run the tapes, Allstate then made a half-hour film which was made available free to any interested groups around the country in late 1972.

Thus, at the beginning of 1973, the air bag controversy continues.

In part, the air bag regulations have become a symbol for both the citizen groups interested in auto safety and the voices urging that new safety changes be introduced gradually after extensive testing and evaluation. The public is trapped between the ardor of the safety groups as they criticize the poor performance of both government and the auto industry in the past, and the dire warnings of the unexpected side effects that may arise if air bags are installed on all new cars.

If we try to reach any sort of a logical opinion from these charges and counter-charges, it is exceedingly difficult to obtain an unbiased evaluation. So little is known as yet about the performance of the relatively new air bag. Furthermore, the past history of the auto industry in safety work is open to severe criticism; even the simple idea of an energy-absorbing steering column (to avoid impaling the driver in a frontal collision) was available for years before federal regulations forced the installation. Certainly one of the great errors in corporate history was the hiring of a private detective to investigate Ralph Nader when he first came into the national spotlight with his attack on the safety of the Corvair in the mid-1960's.

The difficulty the concerned citizen encounters in evolving an opinion on such issues is compounded by the frequent deceptive use of statistics and quantitative data. This ability to present statistics in a desired light is illustrated by a small part of the RECAT report mentioned earlier -- a section in which the argument is made that the federal highway safety program is already quite successful and, hence, we should give the new devices and regulations a chance before requiring the air bag.

This case is a beautiful example of the way statistics can be presented to reinforce a desired point of view. We first look at a small part of the RECAT report (actually the total report is an impressive wealth of valuable information); then we will consider other possible interpretations.

Statement of the RECAT report

Table 6 from page xxix is reproduced below.

Economic Benefits of Reduced Death Rates
for Four Consecutive Years

Year	Death Rate Reductions Relative to 1966 (deaths per 100 million m.v.m.)	Motor-Vehicle Miles Driven (millions)	Lives Saved	Dollar Equiv. Benefit (million)
1967	0.20	961,553	1923	$ 269.2
1968	0.30	1,015,649	3047	426.6
1969	0.47	1,070,600	5032	704.5
1970	0.81	1,121,000	9080	1,271.1

The meaning of the Table is as follows. If we look at the last row, this states first that in 1970 the death rate was 0.81 less than in 1966 (m.v.m. stands for motor vehicle miles). For example, if the 1966 rate was 5.81 deaths for every 100 million vehicle miles, the 1970 figure would be 5.00. The third column gives the miles driven each year; in 1970 this was 1,121,000 million miles.

The fourth column is labelled "Lives saved." Since the death rate is 0.81 less than 1966, there are fewer deaths than would have occurred if the 1966 rate had persisted. Multiplying the 0.81 of column 2 by 11210 from column three gives 9080 fewer deaths in 1970 than would have occurred with the death rate of 1966.

Finally, if each human life is worth $140,000, we obtain the last column: the dollar value of the lives saved. The figure of $140,000 is used throughout the report as the dollar value of one life in order to calculate the benefits and costs of particular safety measures. Whenever we try to measure the total cost of a new health or safety measure, we need to place a value on life in order to measure everything in the same units (in this case, dollars).

The report then goes on to state that: "The estimated benefit realized in 1970 from reduced traffic fatalities was $1.271 billion. Assuming similar declines in injury and property damage rates, the aggregate benefit totalled $3.09 billion. One-third to one-half of this aggregate benefit can be ascribed to the Federal programs for Interstate highway construction and motor vehicle safety regulation.

"Both highway and automobile design improvement programs have resulted in benefits from reduced fatality and injury rates that are greater than the costs attributable to achieving safety, and thus are reducing the cost of automotive transportation."

Report argument

The Report argues that in 1970 the nation saved conservatively $3.09 billion by the reduction in the death and injury rates in the preceding four years. The value may be even greater, since a human life is probably worth more than $140,000. This "improvement" resulted from several factors: for example, stricter enforcement of traffic laws, improved health care, better highways, and the national safety program. The report goes on to claim that, even if only half the savings are attributed to federal highway and safety programs, the dollar savings are greater than the costs.

The implication is clear. The current federal programs are bearing results. Things are getting better each year. Obviously, then, new programs are not needed; all we have to do is continue the existing efforts, and the future is bright. In other words, there is really no basis for the new 1975 measures, and one can not criticize the auto manufacturers or safety program administrators in Washington.

The above paragraphs are a summary of the arguments sometimes advanced by the auto manufacturers and officials in government. They avoid presenting the real problem: the large number of auto deaths and injuries each year. By using carefully selected data presented in such a way as to steer the reader's evaluation to a desired conclusion, the argument creates the appearance of scientific analysis.

This is exactly the mis-use of data which is most dangerous, if we honestly want the public to be sufficiently informed to be able to choose intelligently from among alternatives.

Logical extension of the report argument

If the Report argument has any validity, there's no reason why we should calculate dollar savings only over the last four years. Instead, let's see the "benefit" for 35 years!

In 1936, there were 38,089 deaths and 252,000 million vehicle miles, or a rate of about 15 deaths per hundred million miles. In 1970, the corresponding rate was 5 (one third as much). If we use the Report's argument, somehow in 1970 we "saved" 10 lives for every hundred million miles -- or a total of 112,100 lives in 1970 alone. At $140,000 per life, this represents a saving of $15.7 billion.

Measuring from 1936, we save $15.7 billion the 35th year (1970). If we guess the saving/year increased linearly from zero in 1936 to $15.7 billion in 1970, the total saved over the 35 years is

$$35 \times (\text{Average annual saving}) \text{ or } 35 \times \frac{15.7}{2} \text{ billion}$$

$$\text{or } \$275 \text{ billion}$$

This ridiculous figure comes to $2750 for every licensed driver (enough to purchase every driver a new car) or $1375 per person in the U.S.

It is easy to be convinced that the auto safety program is one of the best buys this country has ever had.

An even better argument goes as follows. The death rate per 100 million miles travelled by horses or horse-drawn vehicles in 1900 was approximately 10 times today's auto rate. Hence, the auto today is saving nine times the lives it takes -- or about 500,000 lives per year. In 1970 alone, the saving was $70 billion (or roughly the total cost of either education or health care in this country).

The ridiculous nature of this argument is apparent if we recognize that, to travel by horse as far as the average American does today by car, everyone would have to spend more than three hours a day riding. Then imagine the urban pollution problem (compared to 1900 when the streets of New York City received daily the relatively minor quantities of 2½ million pounds of manure and 60,000 gallons of urine).

A different view of the same data

The same data of Table 6 of the RECAT report can be presented in a different way:

Year	Deaths	Vehicle Miles Driven (millions)	Increase in Deaths from 1967	Increase from 1967 in Miles Driven (millions)
1967	53,100	961,553	-	-
1968	55,050	1,015,649	1950	54,100
1969	56,210	1,070,600	3110	109,000
1970	55,040	1,121,000	1940	159,400

Now we can argue: The improvement of roads and increase in cars allowed the American people to drive 159,400 more millions of miles in 1970 than in 1967. The cost of this extra mobility was 1940 lives (and perhaps nine times as many injuries).

In other words, the highway-improvement program (for example) gave us greatly increased mobility at the cost of more accidents.

In these terms, a comparison between 1936 and 1970 makes sense. In 1970, we drove 4.5 times as much as in 1936 at a cost of about 17,000 lives. If the highway system of 1936 had remained unimproved for 35 years, the number of vehicle miles driven would certainly not have grown very much. Already in 1936, there were signs of near saturation on many roads. On these highways, an increase in traffic would have reduced the death rate because of lower average vehicle speed.

Comment

In our second view of the same data, we certainly do not emerge with any ringing endorsement of the past efforts in auto safety. Whether an auto safety program is adequate or not depends on

(a) Willingness of the public to sacrifice personal convenience and pay added cost for greater safety.

(b) Readiness of the public to compromise certain aspects of individual liberty (for example, in allowing roadside tests for driving under the influence of alcohol).

(c) Feasibility of new technology to improve safety (e.g., is the air bag technically and economically possible by 1975).

The federal program must weigh such factors and then develop appropriate regulations, subsequently strictly enforced.

Thus, an evaluation of the achievements of the national program so far is to a considerable extent irrelevant to decisions on future policies, except to indicate which policies are likely to result in improvement. Only when the auto death toll is significantly reduced will we be able to argue that the program is, indeed, successful.

Certainly, the most specious argument is that which effectively states that success has been achieved because the death rate per mile is going down, even though in actuality the deaths each year are increasing in number. If one carries this to an extreme, if everyone would drive twice as much next year, we would have a highly successful year if only 90,000 people were killed (compared to this year's 55,000).

3. Alcohol And Driving

Alcoholism is a national health problem, which we might well have discussed in Chapter 3. The relation between serious automobile accidents and alcohol is so startling, however, that the topic has to be included in this chapter. Precise data on the importance of alcoholism are difficult to determine; there is a suspicion that often the easiest explanation for an accident is that one driver was under the influence of alcohol. A study of police records in Massachusetts turned up one such allegation where the driver was a local leader of the WCTU (Women's Christian Temperance Union) who had never had a drink.

Even if the data are viewed skeptically, approximately one half the auto fatalities arise from accidents in which at least one driver is under the influence of alcohol. This seems startling until one realizes that there are 10 million alcoholics and problem drinkers in the United States. In other words, about one licensed driver in ten is in this category. Thus, when you drive, every tenth car coming toward you is, on the average, driven by an alcoholic.[1] Hopefully, many of these are not intoxicated at the time.

Alcoholism in general

Alcoholism is directly related to about 60,000 deaths per year (42,500 of them male -- a significant contributor to the problem of higher male mortality discussed in Chapter 3). The specific causes of death can be further subdivided:

14,500 from directly related disease: cirrhosis of the liver, chronic alcoholism, and so forth

7,000 from suicide and homicide

26,500 from auto accidents

12,000 from other accidents

In spite of the magnitude of this health problem, per capita drinking has not increased significantly if we view the past century, even though the last few years have shown a very sharp rate of increase (Fig. 4-2). The

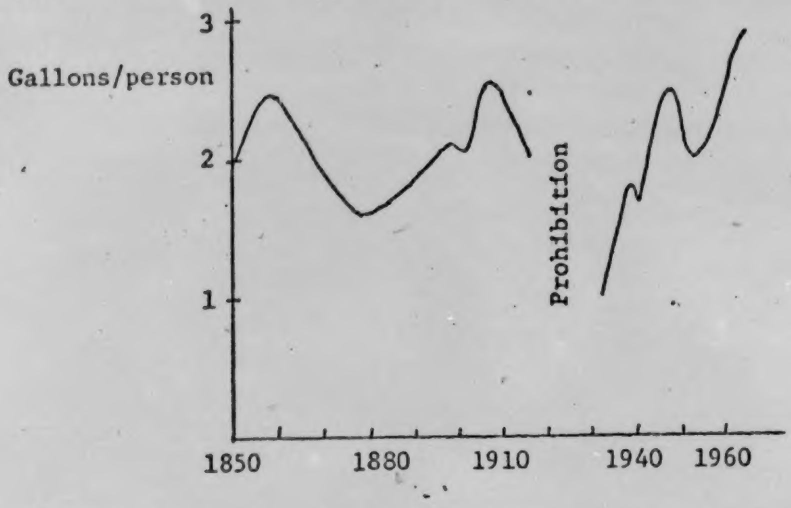


Fig. 4-2 Total consumption per person in the drinking-age population (those over 15) as measured in gallons of absolute alcohol (from wine, beer, and distilled spirits)

situation today is not much different from that in the late 1850's.

---

[1] And the odds are really no better if you are an alcoholic or problem drinker.

One might also measure the seriousness of the alcoholism problem by the annual direct cost in lost work, medical expenses, impaired job efficiency, and accidental damage. This figure has been estimated at ten billion dollars. In the major metropolitan areas, companies frequently contract for permanent accomodations at residential hospitals for their senior-level employees who require a drying-out period.

Alcoholism is not uniform across society. Although a numerically small part of the overall U.S. situation, the 700,000 American Indians pose a particularly serious problem. The average age at death for this group is 44, compared to 65 for the entire U.S. population; 20% have no housing, and 70% of the housing is below any acceptable standard. Some of the schools have a 100% drop-out rate. Under these conditions, it is not surprising that in some tribes every male over 18 has a drinking problem, and it is not unusual to find that the annual number of arrests for drunkenness begins to approach the size of the population. This deplorable status of the Indian exists even though the federal government spends $500 million per year on Indian services -- equal to more than $1,000 for each Indian living on a reservation.

Legal measures

Alcohol is involved in approximately half of the auto fatalities; in the great majority of these cases, the problem drinker is to blame rather than the social drinker (Fig. 4-3). Immediately the question arises: Why not tighten up the laws against driving while under the influence of alcohol?

The author's strongest impression from a visit to Norway arose from

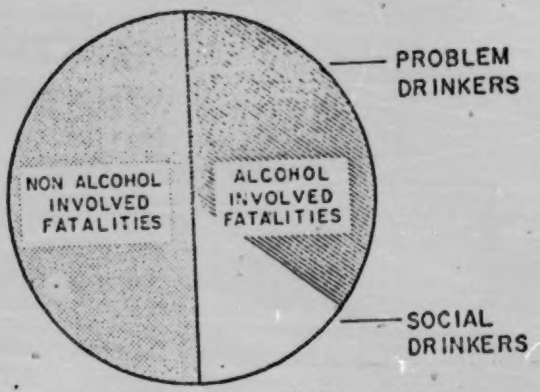


Fig. 4-3 Distribution of auto fatalities

a dinner at which everyone had a social drink or two. After dinner, the host called a taxi to return us to his house rather than drive his own car, because of the very stringent laws in that country.[1] These national actions include

Mandatory chemical tests

Roadblocks to catch drinking drivers

Mandatory three-year license revocation

Three-month jail term for driving while intoxicated

Insurance cancelled if drinking

The result of strict enforcement of this policy is that only 10% of the fatalities (rather than 50%) are alcohol-related.

In this country, in contrast, laws vary among the states and they are often only half-heartedly enforced. The National Highway Traffic

---

[1] In San Salvador, the sentence for drunken driving is death by the firing squad (although one of the author's students, who happened to be from there, emphasized this does not apply to tourists). Needless to say, there are few offenders and no second-offenders.

Safety Administration (NHTSA) of the U.S. Department of Transportation has decided that a tough program is just not possible in this country. Courts will not enforce stringent penalties. The public will not support automatic license revocation or jail sentences and, in response to public attitudes, the courts will allow pleas to lesser charges when the man is accused of driving while intoxicated.

The premise that a tough program against offenders can not be effective in this country seems open to serious question. The legal progress in the last decade has been impressive, and we now find that:

(1) Every state has chemical test legislation, with the level varying from 0.08% to 0.15% blood-alcohol concentration.

(2) Only Illinois does not have an implied-consent law.

(3) Five states now give roadside breath tests.

In spite of this progress, the federal government (represented by NHTSA) takes the official position that the United States can not follow the lead of the Scandinavian countries.

If auto accidents resulting from alcohol can not be reduced by strict laws, what alternatives are available? Two possibilities exist: technology to keep the intoxicated driver off the road, and community or social programs.

Technology

In technology, two different approaches have been proposed to detect the fact that the driver is under the influence of alcohol. The first attempts to lock the ignition when the air exhaled by the driver indicates alcohol.[1] The feasibility of such a device is ruled out by two obvious problems.

---

[1] Time, Jan. 3, 1972, p. 53, reports on a Japanese development.

First, unless the driver's breath is collected directly, other passengers can trigger the ignition lock, or detection can be avoided by opening the window. Second, the lock can be triggered by the fumes from various perfumes and after-shave lotions.

If the public is to accept such devices, the technology must work with reasonable fidelity. There must be a relatively small number of false positives (locking when the driver has not been drinking) and false negatives (failing to lock when the driver has been drinking). This concept of false negatives or positives is a common and important way to measure the trustworthiness of many measuring systems. For instance, in medical tests false positives indicate the disease when it is not present. This causes unnecessary patient worry, and often requires an extensive series of further tests. False negatives mean that the disease has been overlooked, and the chance for early treatment is lost. Thus, the value of a test can be measured in terms of the percentages of false positives and false negatives which are obtained. In the same way, the breath test for the driver will be useful only when the fraction of false results can be kept very small.

There is a second type of technology which has been proposed to prevent driving by people under the influence of alcohol. This approach is represented by a device to display a code number and a dial. A randomly chosen number is flashed before the driver -- for example, 72148. The number then disappears. During the next few seconds, the driver must dial in this same number. If he fails, the ignition will not unlock. Unusally the driver is given a second chance to "play the game," in case he accidentally made a mistake in one digit or was not looking when the original number was displayed.

The principle here is that the driver who is under the influence of alcohol has both a poorer memory and a slower reaction time. Unfortunately, there are again difficulties, and it seems unlikely that such a device will ever be used widely.

(1) Public acceptance is dubious. The average car owner just does not want to bother with such "games" every time he starts his car.

(2) The test could be taken by a sober passenger or bystander.

(3) It is difficult to design any reaction-time measuring system which catches the drinking driver and lets the aged driver pass. Average, human reaction time increases with age (Fig. 4-4).

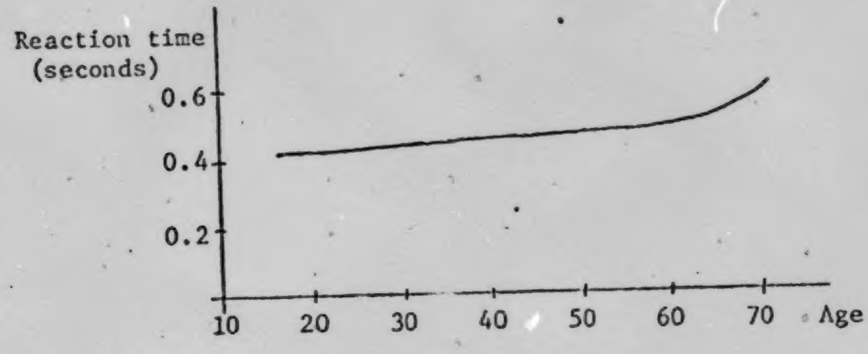


Fig. 4-4 Total human reaction time varies with age, on the average

Reaction time can be measured easily. The subject holds out his thumb and forefinger. The tester holds a ruler with the zero-inch end at the subject's fingers. At any time, the tester drops the ruler. Once he sees the ruler falling, the subject closes his fingers. The distance dropped is a measure of reaction time (4" - 0.15, 8" - 0.2, 12" - 0.24 sec.). This is the minimum reaction time in a simple task; the data of Fig. 4-4 refer to the total reaction time when the foot has to be moved to the brake in a driving emergency.

An interesting variation of the experiment has the subject (a young man) measured completely sober, then again after a bottle or two of beer. For the person not used to alcohol, one beer is enough to slow the reaction sufficiently that he misses the 12-inch ruler altogether. Not only is his reaction time increased, but he is convinced it is as good as ever; he is sure the tester is hurtling the ruler downward.

(4) Any device of reasonable cost could be removed or permanently disabled easily by a technician. (One of the persistent problems with air-pollution-control devices has been the tendency of owners to remove them once the car has passed inspection in order to improve gasoline mileage and increase effective engine power).

Thus, while technology can provide a variety of devices which theoretically can prevent car operation by a driver under the influence of alcohol, all such approaches seem to depend on an unachievable level of public acceptance and willingness to undergo personal inconvenience.

Alcoholic countermeasures program

Confronted by the conviction that legal measures are not a total solution and that technology can not help directly, the National Highway Traffic Safety Administration has started a group of Alcohol Safety Action Projects. In 35 communities across the nation, local programs are totally funded by the federal government at a total cost of $82 million.

Each project is designed by the community and includes a variety of specific steps which, hopefully, will constitute an effective total attack on the problem of alcohol-related accidents.

While the programs vary over a wide range of approaches, the common feature is the hiring of additional justice personnel to locate the problem drinkers and provide individual rehabilitation services. Specific actions include such features as:

(1) For convicted problem drinkers, the option of a jail sentence or a daily use of a drug which causes nausea when alcohol is consumed later.

(2) After a second conviction, the problem drinker is sentenced to 30 days in a residential "hospital."

(3) In San Antonio, Texas, special policemen videotape the behavior of drunken drivers and show the tapes after the arrested man has sobered up.

(4) Special patrols are assigned at night to areas with a large concentration of bars and nightclubs.

(5) In part of North Carolina, the emphasis is on massive public education. Whiskey and wine are sold in paper bags which contain a breath tester and a chart to evaluate breath alcohol concentration.[1]

The long-term hope is that from these local experiments there will emerge a set of tested guidelines to help a community or state develop an

---

1. The proper interpretation of breath tests for determining alcohol in the blood and, then, ability to drive is a fascinating scientific and legal problem. The difficulties, arising because of differing human physiologies as well as the variety of instruments, are described in the article by William S. Lovell, Breath Tests for Determining Alcohol in the Blood, Science, vol. 178, October 20, 1972, pp. 264-272.

optimum alcoholic countermeasures program.

Final comment

The combination of alcohol and driving is a major socio-technological problem. The large number of deaths and injuries are particularly serious because they impact strongly the young and healthy population, and they represent a continuing threat to the large, innocent portion of the public.

What alternatives are available and what can an individual do? The most attractive approaches are not technological, but rather legal and social. Increasingly stringent law enforcement depends on public support. The alcoholic countermeasures program of the U.S. Department of Transportation provides a vehicle through which local efforts can be designed and funded.

4. Emergency Medical Service

In 1971, a soldier injured in battle in South Vietnam was brought to an emergency medical facility in just over 20 minutes, on the average, and he was seen by a physician within 30 minutes of the time of injury. Furthermore, during the interval he was under the care of paramedical personnel who were adequately trained and equipped to handle most of the emergencies.

In that same year, an individual injured in New York state had to wait an average of 69 minutes for care by a physician -- or more than twice as long. The data for New York City (and many other U.S. cities) are comparable. Also, the personnel manning ambulances often had only minimal training, and many ambulances were inadequately equipped.

What are the reasons for the long, 69-minute interval? Actually this is only an average, and the detailed statistics would show many cases much longer. Since a few people are probably injured just outside a hospital emergency room, there must be others for whom the interval is two hours.

In an attempt to understand the reasons for the poor performance of the emergency medical service (EMS), we might try to break down the 69 minutes into its components. The New York State Health Department did this in its massive study of 50,000 ambulance cases during the first six months of 1971. The findings were:

Time until ambulance called after an injury, heart attack or crisis situation	11 minutes
Time for ambulance to reach victim	15 minutes
Time until victim reaches hospital	18 minutes
Time in emergency room before physician sees patient	25 minutes
Total	69 minutes

(The last 25 minutes is probably not as serious as it might seem, since at this point the victim is being observed by nurses. If it were a time emergency, presumably a doctor would be summoned more rapidly).

This breakdown into the four categories places in evidence the weaknesses of the EMS system in many parts of this country. Discussion of EMS is diffult because performance differs so widely from one region or town to the next (usually without citizen awareness of the quality of emergency care available in his own area). In a few towns, there are major efforts to develop a logical, unified system using modern communication and transportation technology and including the best possible training of personnel. In other cities, the EMS system is a hodge-podge of jurisdictional disputes, with elements of care provided by hospitals, police, fire departments, and private ambulance services -- groups which may compete rather than cooperate.

In an attempt to bring order and progress to this often chaotic situation, the federal government (through the Department of Health, Education and Welfare) initiated pilot projects in 1971 -- five local efforts to demonstrate what could be accomplished through a centralized regional program adequately funded. The improvement of emergency services was adopted as a key element in the Administration's program for the use of science and technology for the benefit of the public. Expansion of the effort in 1972 became temporarily the victim of the attempts to hold down federal expenditures, but there is every anticipation of continuing and increased emphasis on this area during the next decade.

In the remaining paragraphs of this section, we want to look at each of the four time periods mentioned above; our objective is to understand ways in which the total time can be reduced and the emergency medical care improved.

Calling for help

In many parts of the country, one still has to dial a seven-digit number to reach the source of emergency care. If the call is from a pay phone, a dime is still needed, even to reach the operator. If the call is to the police, there may be a very limited ambulance and emergency service available and under police control. There is likely to be no regional control center where information on availability and location of all ambulances is kept up-to-date, so that the nearest, properly equipped vehicle can be dispatched to the scene.

Two relatively simple technological improvements are obvious:

(1) Pay phones should be designed so that the operator or emergency numbers can be reached without requiring the caller have change to actuate the dial tone. This system, already in operation in the United Kingdom, is gradually being installed in this country.

(2) The emergency number should be uniform throughout the United States. In the Soviet Union, one can dial the same two digits on any phone. The comparable approach in this country is the 911 number to reach the police, although as yet this is by no means nationally available.

The importance of a central source for help and information has been demonstrated by the poison control centers which have been in operation for several years in different parts of the country. If a child swallows a substance feared poisonous, the mother can call the center, describe the substance, and immediately receive first-aid advice (the operator at the center has a catalog of information on appropriate antiaotes). The centers also yield important information for new regulations to decrease the number of serious poisonings. For example, early experience with the centers revealed the large number of cases in which very small children ate all the baby aspirin tablets in a bottle, a step encouraged by the candy flavoring added. As a result, the number of baby aspirin which can now be sold in one container is limited so that severe poisoning can not occur. Such an imaginative, even though trivial, decision represents a small step forward in consumer protection, but a step which probably would never have been conceived without the information input from the poison control centers.

Help to the scene

The second time interval is that required for the ambulance or emergency vehicle to reach the victim. The obvious problems here are the location of waiting ambulances and the techniques used for dispatching. In New York City, merely as an example of a large urban area, there are 109 ambulances stationed at 49 hospitals throughout the region. Calls to the police lead to instructions to the nearest hospital to dispatch an ambulance.

In London there are now attempts to establish a centralized system. When a need for help is perceived, the central computer will determine which of those ambulances appropriately equipped for this call is closest in time to the scene. Once the selected ambulance is then called, it will be routed through the city streets by radio communication. Ultimately such a system might also keep a running record of which blocks are severely congested in order to route the emergency vehicle around bottlenecks. The total system requires classification of the medical equipment carried by each vehicle: if the call is for an apparent heart attack, the answering ambulance will have coronary care capabilities and correspondingly trained personnel.

One aspect of the system which until recently has received very little attention is the desirable design of the warning devices (lights and sirens). The evaluation of a siren-light combination turns out to be exceedingly difficult. The purpose of the equipment is to alert drivers and pedestrians to the approach of the emergency vehicle. This warning must be given even in the presence of a distracting environment (flashing neon signs and horns). It must be effective even when the driver has his windows closed and radio on. It must work with drivers who are tired, partially deaf or color blind. Finally, it should minimize the disturbance to nearby residents and people behind or well to the side of the emergency vehicle.

In the light of these requirements, current research is focussed on both new technology and people-technology interactions. In the former category, sirens and lights are sought with improved directional characteristics, so that the warning is focussed on the street ahead and entering side streets. A signal will be transmitted from the emergency vehicle to control the traffic lights a block ahead in order to stop traffic and pedestrians from entering the street in the path of the vehicle. In the latter category, research is studying which types of sounds and lights are most successful in alerting the human being.

Transporting the victim

Once the ambulance reaches the patient, there is the problem of transportation to the hospital. Again, the routing problem arises. It is not unusual for an ambulance to take the patient first to a hospital which can not accept him; this even happened after Senator Robert Kennedy was fatally shot in Los Angeles in 1968.

Ambulances are often poorly designed and equipped. 80% of the 25,000 ambulances in the U.S. are converted hearses, limousines, or station wagons. Only one third of ambulance attendants have received the 80-hour training program which is considered minimal. Furthermore, there is often no direct communication link between the vehicle and the hospital, so that trained physicians can advise on the care during the travel period.

At the emergency room

Finally, many hospital emergency rooms are now inundated with ambulatory patients or out-patients. For a large fraction of the population, the emergency room represents the principal means of access to the health care system. From the standpoint of emergency medical care, the ideal system (even if it is too costly) would be one in which the emergency room was staffed by doctors especially trained for trauma treatment and was used only for its primary purpose.

Final comment

The system for emergency medical services in the United States is one in which local control is the custom. Consequently, an enormous variation in the quality of service has developed, dependent on the unity and centrality of local planning, the financial support given by the community, and the demands by the public for performance. As confirmed in recent studies by the National Academy of Sciences, the Department of Health, Education and Welfare, and the White House Office of Science and Technology, the result is on the average an emergency care system well below the standards in other nations (the United Kingdom and the USSR). In many cases, improvement is blocked by the difficulty of bringing together the various groups which are already involved: the local medical society, police, firemen, hospital personnel, and so forth.

5. Other Accidents

Automobiles account for only about half of the fatalities and serious injuries. In the preceding sections of this chapter, we have focussed on auto accidents simply because this is the area which has received the most public attention in recent years. It is also the direction in which it seems that it should be possible to make the most improvement with relatively straightforward applications of technology, new legal and regulatory measures, and enhanced public awareness and concern.

Other areas could be considered. Fire causes more than $5 billion annual damage in the United States. While there is still great ignorance of how fires start, spread, and can be controlled, there is a variety of programs which hold promise of decreasing both fatalities and property losses. A few of these serve as an indication of the promise of this research:

(1) Just the past few years have witnessed the first studies (including computer simulations) to determine the optimum location of fire stations and equipment within a town to minimize the expected time for the equipment to respond to a call. Working for East Lansing, Michigan, a Bureau of Standards group has developed an optimization procedure.

(2) In most major cities, false alarms constitute about 50% of the calls.[1] The New York City Rand Institute showed that the corresponding loss in performance could be cut significantly by using cruising cars to investigate alarms and turn back the major equipment if the alarm is false.

(3) Once the firemen have reached the scene, fire fighting effectiveness can be improved by using hydrants which can be turned on easily by firemen (but not, of course, by youngsters on a hot, summer afternoon), slippery water (the water with a polymer chemical added to ease flow along the hoses and increase pressure at the nozzle), and robots to carry hoses

---

1. One effective way to reduce false alarms from call boxes is to add picturephone service so the caller is visible to the operator. The problem is the maintenance of the picturephone equipment. When in some cities the half life of a pay phone on the street is only a few hours, picturephones are obviously susceptible to vandalism.

and equipment into buildings when there is danger of the structure collapsing.

(4) The National Bureau of Standards is now trying to develop safer matches and lighters. The primary goal is to make it nearly impossible for a child to ignite them, since almost half of the victims of accidental fires caused by match ignition are children under 10, while children under 6 account for 29%.

(5) In the same vein, the Department of Commerce is responsible for setting the standards under the federal Flammable Fabrics program. In July, 1972 the first standard was issued for children's sleepwear up to size 6 in spite of a vigorous campaign by segments of the apparel industry to water down its provision. An attempt was made to phase in the standard by requiring immediately that all garments which do not meet the standard through 50 washings must carry a label warning the purchaser to keep the clothing away from high heat. After 1973, all apparel sold must meet the standards.[1]

(6) In many cities a major effort is underway to determine appropriate building codes to control the effects of fires in modern skyscrapers. Two serious problems have appeared in recent years. The first difficulty arises because smoke tends to fill the stairwells from the floor on which the fire is located. The occupants are then blocked from using the stairs for emergency exit and smoke can be carried to other floors. The answer

---

[1] In England in the late 1960's a rigid standard was introduced abruptly. Prices of infant nightclothes rose so rapidly, many mothers bought material (which usually did not meet the standard) and made the clothes at home. The result was no real improvement in child injuries and fatalities.

now under study is to keep the air pressure in the wells slightly above atmospheric. If the door to the fire area is then opened, the flow of air is from the stairwell into that floor.

A second problem arose with the elevators. In some buildings, heat-sensitive call buttons were used: the heat from the person's finger actuated the button circuit to call the elevator to that floor. The device was attractive because there was no mechanical switch to close (and hence to wear out). Unfortunately, with a hot fire on a particular floor, all elevators were automatically called to that floor. The obvious answer here is to return to mechanical switches or buttons actuated by the change in electrical capacitance when a person's finger comes near. Furthermore, computer-controlled elevators should include special programs which take over as soon as a fire alarm is given -- programs which reserve certain elevators for fire-fighting personnel and schedule others to assist in building evacuation.

These specific examples show the range of current, engineering research designed to reduce the human and economic costs of fires. Analogous programs are underway in such areas as injurious side effects of chemicals and drugs, both those administered directly to people and those which ultimately reach people after use with animals, as fertilizers or insecticides.

6. Conclusion

The most interesting feature of this field of emergencies or accidents is that the major effort by so many federal agencies to cut down the human damages is a remarkably recent phenomenon. The Washington program is illustrated by these activities:

Auto safety program -- Department of Transportation

Emergency medical services
Chemical accidents } -- Department of Health, Education and Welfare

Fire research -- Department of Commerce

Reduction of earthquake damage -- National Science Foundation

Control of radiation damage -- Atomic Energy Commission

Most of the other cabinet departments and independent agencies have major efforts in the reduction of accidental injuries, deaths, and property losses.

Yet all of these programs have either been initiated or at least gained major status only in the last five years. The emphasis has grown from Ralph Nader's crusade and the entire consumerism and environmental movements. To a scientist or engineer, the most remarkable manifestation is the totally changed, Washington attitude which undoubtedly reflects public opinion.

Less than a decade ago, science and technology were handsomely supported by the federal government primarily because science was inherently desirable and good. Today there is a widespread opinion that science and technology must be selectively supported to ensure emphasis on developments which may improve the quality of life for everyone. Technology is good insofar as it serves the public purposes; science must find an appropriate balance between augmenting knowledge for the sake of the beauty of knowledge alone, and contributing to the amelioration of major national problems.

As with other major social movements, one can trace the consumerism emphasis far back in history. About 2050 B.C., Ur-Nammu (founder of the

Third Dynasty of Ur) introduced official weights and measures to restrain dishonest merchants. By the Roman Empire, the motto "Caveat emptor" or "let the buyer beware" was generally accepted.

What is new in the last five years is the location of this movement in a central position on the national scene. By 1972, 43 of the states had governmental offices to aid the consumer; 39 federal agencies and departments had specific responsibilities and programs. The public was bombarded by relevant news stories in the press and on television and radio. This change obviously resulted from a variety of social, political, and economic developments. But any study of causes would have to attribute major contributions to Nader and his associates and to the forces brought to bear on the government by the concerned young people and minority groups. The national emphasis on the critical importance of socio-technological problems has led to an insistence that science policy (as well as economic policy) accept a responsibility to contribute to the quality of life.

Yet the problems are not trivial. Even the simple, clearly defined problems often seem to have no obvious solution. One final example, drawn from the subject of accidents which is the theme of this chapter, illustrates the difficulty.

The electric, pop-up toaster has heating coils surrounding the bread. If a slice of bread is stuck in the toaster, the housewife commonly uses a knife or fork to try to pull it out. If the toaster happens to be still on, there is a serious danger she will be electrocuted. While the number of people killed is not large, this seems like a totally unnecessary risk.

The technological solution is easy. The heating coils are simply encased in a substance which is an electrical insulator (just as the coils on a range). Unfortunately, there are two side effects. First, the new model is slightly more expensive than the conventional toaster. Second, because the insulating material also hinders the transmission of heat (by a basic law of physics), the new model is slower toasting.

Recognizing these limitations, the appropriate government regulatory agency really has only one option. The only way to ensure that the new toasters are used is to require the safety feature. Such an action would raise a variety of questions:

(1) Is this carrying the role of government too far? Are we using government to try to protect everyone to a ridiculous extent?

(2) Does the increased cost of the new model mean that more of the economically disadvantaged are now deprived of the convenience of a toaster?

(3) If toaster sales are significantly reduced, what happens to employment in the industry?

(4) Can the public be educated or informed about the situation?

(5) If toaster use drops, will people find a new way to accomplish the same goal; if so, are even more accidents then likely?

While the toaster example is certainly not very significant unless you happen to have a friend or relative who has been a victim of such an accident, it does point out the difficulties we encounter as we attempt to use technology to decrease the frequency of accidental injury and death.

CHAPTER 5

DECISIONS

Will abortion be legal? Will a new airport be built a mile offshore? While these two questions are certainly unrelated in substance, in several states they are both strong political issues. The abortion question often is viewed as a strictly moral and ethical question, although a different part of the electorate emphasizes the aspects based on social science and medical science. The airport question, in contrast, is likely to depend on environmental science and economics.

The decision in either case (and certainly in the airport problem) is likely to be based on the public understanding of the scientific knowledge related to the issue. Certainly public opinion can be strongly directed by emotional and political arguments, which may even distort the scientific knowledge. Furthermore, certain questions may be so strongly moral that the scientific elements are unimportant. The basic article of faith on which this book is based, however, is that in many cases the decision is likely to be better, the more understanding the public has of the scientific issues.

The ultimate goal of technological literacy is better decisions. If we consider the airport-location problem, technology can not solve the problem--it can not provide an economic, convenient location which does no damage to the total environment. Technology can present to the public and its political process a picture of the options--a model of the problem which includes the advantages and disadvantages of each alternative.

Furthermore, technology can present some understanding of the decision process. Once the problem is understood, how can the decision be made? In this chapter, we want to look at a few of the characteristics of modern decision techniques. Given an understanding of the problem, how do we make an optimum decision?

1. Drug Control as a Decision Problem

The quality of life in a large city depends on a host of different factors. What fraction of the people is on welfare? Are jobs increasing? Are schools good? Is health care available for most of the people? Is transportation cheap and convenient? What is the air quality?

These factors are inter-related. When we take an action designed to improve one aspect, we are likely to have a major impact on others. In this sense, the city is a system--indeed one which is incredibly more complex than any system built for the space program or in industrial automation.

This city system is like an inflated balloon. When we push in at one point, it bulges out somewhere else. If we improve mass transportation (for example), we are likely to find taxes rising, then more industry leaving the city, and finally a more severe welfare and unemployment problem. Better commuter transportation encourages the middle class to move to the suburbs; in the central city, housing deteriorates, crime increases, tax income falls, and so forth.

This example alone leads to a very pessimistic attitude: with any urban problems, maybe the best approach is to do nothing. Possibly the cure will be worse than the disease.

Naturally, no one can accept such a negative attitude. There is a problem, however, in finding the right way to impact a city system. What action or combination of actions will yield a net positive benefit?

Counter-intuitive behavior

Jay Forrester has used the adjective "counter-intuitive" to describe complex systems in which the result of an action is exactly the opposite of what one would expect from simple or intuitive reasoning. For example, a principal public effect of widespread drug addition is crime, especially burglary and robbery. In order to reduce the crime rate, the government launches a major program to decrease the supply of heroin. As a result, the price of heroin rises sharply, and more addicts turn to more serious robberies. The net effect is counter-intuitive: an anti-drug program leads to increased crime. Furthermore, the higher heroin prices attract into the business more experienced criminals. With the incentives of greater profits, criminal efforts to spread the habit among teenagers are intensified.

Thus, a stronger anti-drug campaign can lead to more crime and a larger number of addicts. One might suppose, then, that the undesirable effects could be reduced by the opposite approach: easier access to drugs. The British tried this in the late 1960's, indeed even giving drugs to addicts. The result was a rapid increase in the addicted population, simply because of easier access. From the economics of the criminal side, when the price is low and profits are small, a large market must be created to maintain total profits for the distributors.

Finding the optimum

Just the arguments above show that either the crime rate or the number of addicts depends on the strength of the government anti-drug program in a complex way--perhaps as shown in Fig. 5-1. There is an optimum point at which we should work. The government control program should be maintained at level X. Any change from this level in either direction results in a worse situation. We can not reduce the addict population to zero (any more than we were able to stop drinking during Prohibition); we have to try to locate the best possible government program.

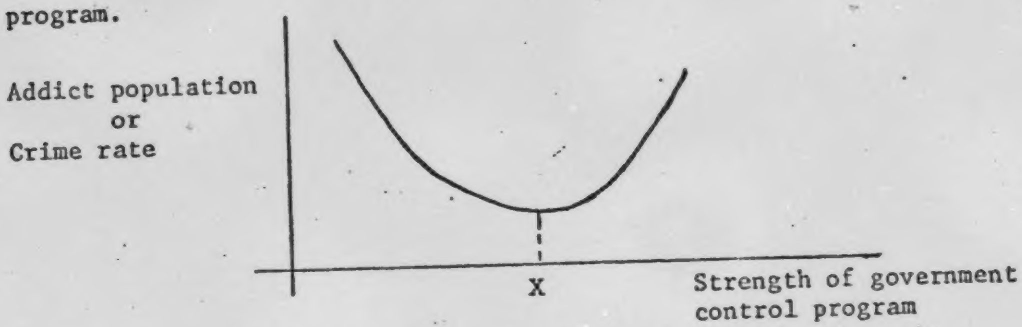


Fig. 5-1 Number of addicts as it depends on the intensity of the control program.

The problem is even worse than described above. At any given time, we do not know the shape of the curve in Fig. 5-1. All we know is that we have a certain strength or intensity of the government control program, and that the addict population seems to be at about a certain level. We never can measure the addict population with any real accuracy or confidence. Thus, we only know that we are located somewhere along the line segment of Fig. 5-2.

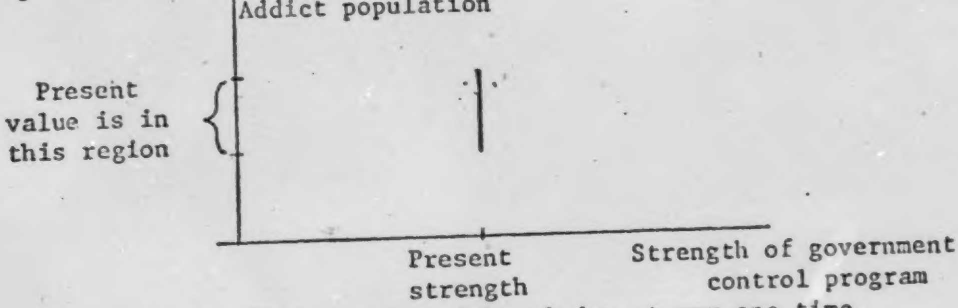


Fig. 5-2 The extent of knowledge at any one time.

If our goal is to decrease the size of the addict population, we need to know whether we are to the right or the left of the minimum of Fig. 5-1. In which direction should we change the strength of the government control program to cause a smaller addict population?

The first thought might be to increase slightly the control program, then measure the new addict population. If there were a decrease, we would deduce that we originally were to the left of the minimum in Fig. 5-1--that the control program was too weak. By intensifying the control program step-by-step, we can move steadily toward operation at the minimum point.

This search for the optimum is illustrated in Fig. 5-3. Here we start at A. We intensify the control program to B and the population drops so we are at C. We again intensify the program to D, the population drops to E. This "search" technique continues until we pass the minimum or optimum. Then a further intensifying of the control program causes the population to rise. As soon as this effect is observed, we back up and rest at (or very near) the best possible condition.

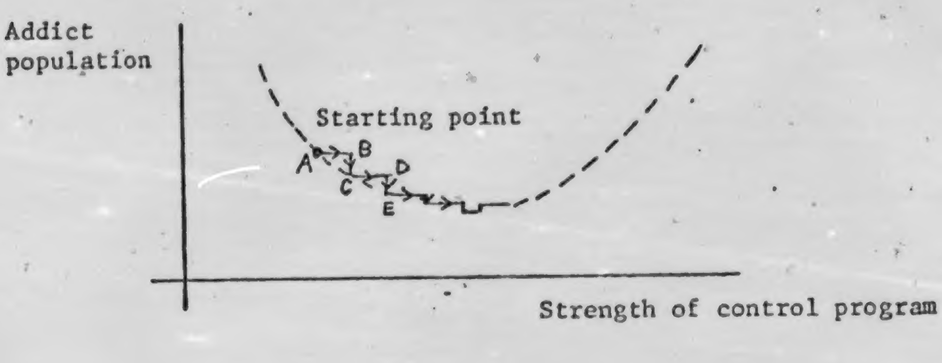


Fig. 5-3 Search for the optimum.

This search technique for finding the optimum is a fundamental concept of modern technology. The human being uses this hunt-and-try approach in an enormous variety of tasks. When he goes to bed to sleep, he searches for a maximally comfortable position; driving an automobile on the open road, he tries various speeds until he homes in on one which results in a desirable compromise between driving comfort and desire to minimize travel time.[1]

Unfortunately, in the addict-population control problem, the method is not very useful, primarily because of the fuzzy model shown in Fig. 5-2. We just can not measure the population size with any accuracy or confidence. Consequently, when the control program is intensified (Fig. 5-4), it is not at all clear whether the population has increased or decreased.

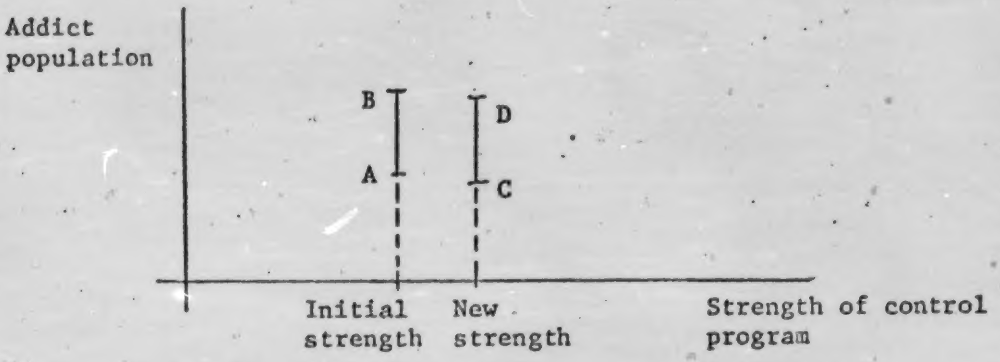


Fig. 5-4 The problem with fuzzy measurements.

---

[1] This search technique is the basis for many of the systems designed recently for the space program or for industrial automation. The system is built to search automatically for the best operating point. The fact this approach works in such cases and fails in the drug control problem is a result of the much greater complexity of urban or social systems.

Initially, we are located somewhere between A and B. The control program is intensified. Six months later, attempts to measure the population show we are somewhere between C and D. The two bands are so wide, the uncertainties so great that we just can not draw any conclusion. Furthermore, in that six-month period, other social and economic changes have occurred which may well have totally overshadowed the effects of the change we put into the system.

Thus, we are faced with a problem in which it is probably impossible to find a best decision.

2. Complexity of the Drug-Control Problem

As described above, the drug control problem is enormously simplified. We have assumed there is only one way to impact the problem: we can increase or decrease the strength of the government control program. Similarly, we have used only one measure of the seriousness of the problem: the addict population. Even this seemingly trivial problem appears to be insolvable.

In actuality, the problem is enormously more complex. Even if we restrict our attention to heroin, there are many different parts of a government control program. A decision on an optimum program has to include descriptions of what effort to put into each of the various directions. Furthermore, the seriousness of the problem is also multi-dimensional (if we use the terminology of modern technology, where we simply mean that there are many different measures which must be considered to obtain an overall picture of the magnitude of the problem).

First, with respect to the various dimensions (or parts) of the control program, the government tries to impact three separate parts of the system:

(1) Manufacture

(2) Distribution

(3) Addict treatment

In the area of manufacturing, there has been considerable publicity given recently to the U.S. attempts to reduce the poppy growth throughout the world. The U.S. relations with Turkey, northern Africa, and southeastern Asia have been strongly influenced by the problem. (Burma, Thailand, and Laos represent the major opium production center). With the improved cameras now used for satellite observation of the earth's surface, it is possible to determine the locations of poppy growth.[1] The proposal has been made in the press that the U.S. unilaterally should eradicate the poppy crop in any nation which refuses to control growth, through use of selective herbicides and biological methods.

There are two obvious weaknesses in any proposal of this sort. First, what right does the U.S. or any group of nations have to control the agricultural activity of another country? Second, and perhaps even more important, the prices paid for heroin in the U.S. are significantly higher than anywhere else in the world.[2] Consequently, a shortage of heroin will have little effect on the U.S. supply unless the scarcity is acute. It has been estimated that 95% of the world's poppy crop would have to be destroyed before there was a significant decrease in the U.S. availability of heroin. The poppies can be grown in relatively small and isolated fields; only 10 square miles are required to supply the U.S. addict population of 560,000. Thus, crop eradication seems totally unpromising unless one takes the global view that it would be desirable to reduce the worldwide supply of heroin.

---

[1] Different agricultural products give different photographic signatures--that is, different combinations of color and heat radiation.

[2] The U.S. addicts consume between 7 and 10 tons of heroin a year, an amount made from at most 100 tons of opium--a small fraction of the 2500 tons produced annually in the world.

The manufacture phase encompasses not only the growing of poppies, but also the processing of opium to create the heroin. Again, it is extremely difficult, if not impossible practically, to impact this system, with its many small, low-capital businesses scattered so widely over the world and with the extraordinary profits associated with sales to the United States.

Because of the publicity given to the manufacturing in Marseilles, France, a Long Island group compaigned during 1972 for a public boycott of imported, French wines until the French government closed down the processors in that country. Even if such a boycott were effective, there is no indication that there would be any significant change in the U.S. supply of heroin (although there might be benefits for the New York and California wine growers--one would imagine they would be enthusiastic supporters of the boycott).

If the manufacturing part of the system can not be significantly affected within moral and diplomatic constraints, perhaps the distribution network can be. The first thought is to close the source of supply at the point of entry into the country--specifically at customs inspection. Compared to most of the other advanced countries of the world, the U.S. already subjects entering travelers to an unusually thorough inspection of luggage and packages. To detect the smuggling of heroin in modest quantities would require careful searches of the persons of all entrants.

Such a procedure would certainly evoke loud complaints from individuals concerned with the inconvenience, long time delays in clearing customs, and infringement of individual rights. In 1972, the very cursory examination of carry-on luggage at airports in the hope of deterring potential airplane hijackers led to a series of court actions arguing that the government or airline has no right to search personal belongings without the consent of the owner. (Incidentally, these inspections did also uncover numerous cases where air travelers were carrying illegal drugs, even though checked baggage was not opened and any illegal materials could have been shipped there).

Affecting the distribution system after the heroin is in the country is equally difficult. Even when a distributor or retailer is arrested, he is often free on bail within a day or two at the most. The Japanese have taken the approach that the arrest can only be effective if the man is kept off the streets long enough so that the distribution chain is broken longer than the addict can tolerate. Consequently, they use a special Emperor's arrest, which allows detention for four weeks without bail. A similar procedure in this country would certainly generate ferment and solid objections from those deeply concerned about our philosophy that a man is assumed innocent until actually proven guilty.

A basic problem in impacting distribution is that so little apparently is known about the detailed structure of the system. In order to decrease the signal flowing through a system, we need to understand the various routes from source to consumption and the fraction of the flow along each of the alternative paths. Otherwise, we are likely to exert major effort to close a path which is actually minor, or which can easily be bypassed.

One proposal to learn more about the distribution network is to allow known manufacturing plants to continue operating, since shutting them would not decrease supplies much anyway. At these factories, distinctive radioactive tracers would be added to the heroin. Then, when a supply with this tracer is found in the U.S., we would at least know the original source. Hopefully, over the years, a picture could be constructed of the distribution network.

Finally, the government can also attempt to influence addict treatment, or more generally the population of addicts. In addition to the various rehabilitation and maintenance programs, such as the methadone effort, we would include here the education programs now becoming so commonplace in the schools.

Evaluation of these education programs is no easier than many other educational endeavors, and serious questions have been raised as to the usefulness of school courses. There is a trend toward moving such efforts from the high schools into the junior high schools and the upper elementary grades (drugs are the major cause of death in New York City between the ages of 15 and 30, so the emphasis is on starting the educational effort when the student is entering his teens).

Even the very brief discussion of the drug-control problem in the preceding paragraphs is perhaps sufficient to emphasize the awesome difficulty of evolving a rational federal policy. We have not mentioned the major sociological and psychological problems: what factors in the environment or the individual personality place certain young people in a high-risk category? What social, educational, and home life characteristics can be strengthened to reduce the number of young people using drugs?

In broader terms, this country faces the fundamental question: Is drug use an epidemic which can be eradicated, or is it a natural consequence of our present way of life? If the latter, does this mean that we have to accept a certain level of use and addiction in the foreseeable future? If so, our efforts should be directed toward determination of that minimum level and the national and local programs associated with it.

These first two sections have discussed certain aspects of the drug-control program in order to emphasize three characteristics which are associated with so many of the decision problems discussed in this book:

(1) The existence of only the vaguest and most inadequate models

(2) The importance of probabilistic situations--we can only speak in terms of probability; we seldom work with situations in which a certain policy can have one and only one result

(3) The fact certain problems may just not be solvable, at least not the way we would wish.

3. Importance of Information

In the drug-control decision problem, possibly the dominant feature is the nearly total lack of reliable, quantitative information for the model. Very, very little is really known about the system. Any quantitative model must include many estimates or guesses about particular parameters.

The fact that the model is so susceptible to error always raises the question: is there any real value in using the model as a basis for decisions? Since we admittedly know so little about the system, might we not make better decisions merely by using our intuition or our "feel" for the problem? Does an attempt to be quantitative really help when the problem is so poorly defined?

There is no simple answer to these questions. There obviously are problems in which the hunch or intuition of the person who has to make the decision is at least as good as a more logical or analytical approach. Man is remarkably good at making decisions just by thinking about a problem long enough.

There are, however, many decisions in which the more information we can obtain, the better the decision. As the amount of relevant data grows, it becomes more and more difficult for a human being to make a logical decision. Then he has to turn to mathematics or analysis; if the data are sufficiently plentiful, a computer becomes essential.

This concept of "the better the decision, the more, relevant data upon which it is based" is illustrated by a famous problem in decision theory.[1] We are all directors of the U.S. Widget Company, the manufacturer of an outstanding line of widgets. As often happens with a Board of Directors, we have no idea what widgets are or how they are made. We do know that the company can produce exactly 200 each day. Furthermore, the widgets are made in three different colors: red, white, or blue. Because the manufacturing process is so complex, each day the company can make only one color. That is, all 200 widgets made on any one day must be the same color.

Lately, the top management and Board of Directors have been disappointed by the lack of growth in sales. We believe that business and economic growth is an essential characteristic of a healthy organization; otherwise, new people can be hired only when someone leaves or retires, and the organization tends to be increasingly dominated by old men. To stimulate growth, we on the Board have just authorized the sales force to launch a campaign based on the promise: "Double your money back if we don't deliver an order within 24 hours." Obviously this is a dangerous campaign; if the company is too successful, it will go bankrupt simply because of the limited production capacity. But desperate measures are needed.

---

[1] A fuller discussion is given by Myron Tribus in the IEEE Transactions on Systems Science and Cybernetics, SSC-4, 3, 9/68, 241-8. Between service as engineering dean at Dartmouth and Xerox vice-president, Tribus was Assistant Secretary of Commerce for Science and Technology. In that capacity, he made important contributions in his attempts to apply decision theory and engineering concepts to important socio-technological problems--e.g., the establishing of flammability standards for infant nightwear.

When the Board approved this sales campaign, it discussed at length the problem imposed by the limitation that widgets of only one color could be manufactured each day. It was obvious to the Board that the decision each morning as to what color to make that day would be critical to the company's success. The color manufactured must be that which will minimize the possible loss resulting from the "double your money back" guarantee. With customary modesty, the Board decided only it could make the decision. Consequently, at 9 a.m. each morning, the Board meets to vote on which color to make that day.

We now have completed the background for the decision problem. The Board convenes the first morning to vote. The President of the company has anticipated that the Board members will have no idea how to vote unless they have some information. Accordingly, at 8:45 he called the warehouse and asked how many widgets of each color were then in stock. The numbers he reported to the Board are:

	Red	White	Blue
Number in stock	100	150	50

The Board met and learned the above data. The discussion was brief (there really wasn't much to talk about). Everyone who voted selected Blue.

Several members refused to vote; they claimed they didn't have sufficient information to make an intelligent decision. Actually, by refusing to vote, they really were saying that they would leave this critical decision to the other members; presumably the non-voters felt that they were less able to make an intelligent decision than their colleagues.

During the discussion, several Board members demanded that they be provided with more information. It was finally agreed that, as soon as possible, the Board would be told not only the current inventory, but also the average daily selling rate over the past six months. A few days later, the following information was presented to the Board:

	Red	White	Blue
Number in stock	100	150	50
Past average daily selling rate	50	100	10

Now the vote was no longer unanimous. The majority selected White (which indeed is mathematically the best or optimum decision on the basis of these data only). Several, however, still voted for Blue or favored Red.[1]

The first line is the same as in the first vote. Does the fact that people changed their minds because of knowing the additional data mean that the earlier decision for Blue was wrong? Of course not. It simply means that additional data allow a more intelligent decision. As the Board members learn details of the model which describes the system, they are able to make better decisions. One has to make the best possible decision on the basis of the available information.

---

[1] This entire discussion of the widgets problem can be more effectively presented in a group or class discussion. The group serves as the Board of Directors. After each vote, a spokesman or two from each element is selected to present arguments for his vote. After these arguments and discussion, a revote is taken before the next stage of data is presented.

Impressed by the more knowledgeable decision they were able to make with more data, the Board then requested any additional information the management could provide. Several days later, the following model was presented

	Red	White	Blue
Number in stock	100	150	50
Past average daily selling rate	50	100	10
Average lot size	75	20	20

For each color, the average lot size was determined--that is, the total number ordered over the past six months divided by the number of orders. When people buy Red widgets, they tend to order large numbers. With Red, there are on the average only two orders every three days, whereas with White we can anticipate five orders a day.

Confronted with these data, several of the Board members pondered for a long time. One member expressed the wish that he knew more about the statistics of the system. For example, how often can we expect two orders for Red the same day; how often do a day's orders exceed the number now in stock.

When the vote was finally taken, the 24 directors divided 9 for Red, 9 for White, 3 for Blue, and 3 refusing to vote. After efforts to convince the three abstainers to vote failed, brief arguments were heard for both Red and White. Then the Directors were asked to choose between these two alternatives only. The winner was Red. (The result again is actually in agreement with the optimum obtained mathematically; the human being is quite good at optimization).

As the directors were preparing to adjourn, the phone rang. It was the star salesman calling to report that last night he had been successful in landing an order for 50 Blue widgets. Since these were now sold and the day was just starting, the data above really should be changed:

	Red	White	Blue
Number in stock	100	150	50
Past average daily selling rate	50	100	10
Average lot size	75	20	20
Number sold before day begins	0	0	50

The Chairman of the Board hurriedly called for a revote. What do you suppose the result was now? What would it have been if the last sale had been 60 Blue? 40 Blue?

Commentary

The interesting feature of this example is the way the decision changes as more data are collected. At each stage, the decision made is correct; it is optimum. As we know more, however, we can make a more learned or more intelligent decision. But even the smallest quantity of data (the number of each color now in stock) permits a better decision to be made than with no data available at all.

The fact that the new data at each stage changes the decision means only that these new data are themselves sufficiently improbable that they do change the structure of the problem. When we make the first decision on the basis of only the number in stock, we assume that the average daily selling rates are essentially the same for the three colors. To change the first decision, these rates have to be markedly different (as they actually are when we have selected the numbers for the second row).

Thus, the argument of the decision theorist is that one should make the optimum decision on the basis of as much information as can be obtained. In broader terms, the cost of collecting more data must be weighed against the increase in the value of the decision resulting from that better data. Even in the widgets problem, the Board can demand so much data that extra people have to be hired just to collect information. At some point the cost of data collection exceeds the added value.

Furthermore, the example demonstrates the increase in the difficulty of decision-making as information increases. The simplistic solution is always easy to find; one needs to know practically nothing, and one's mind is uncluttered by "confusing," contradictory facts. As we learn more about significant decision problems, however, it becomes clear that there is no obvious answer. We always have to strike the best possible compromise among the conflicting demands and pressures.

Even in the simple widgets problem, a slight increase in information would move the problem beyond the limits of ordinary human understanding. For instance, if for rows 2 and 3 we give not only the averages but also the probabilities of different amounts, the problem is so complicated that a computer is desirable to find the best decision. Then we must also decide firmly on what we wish to optimize: do we want to maximize the expected company profit for the day, or to minimize the possible losses from inability to fill orders? In other terms, we must select the criterion for determining the best decision.

The widgets problem epitomizes an important object of this book--the demonstration that the wisdom of decisions increases as the understanding of the system becomes more complete.

4. Simple Choices

The easiest decisions are those from among a small number of choices. In this section, we cite two examples of such decision problems. Each of these depends strongly on the concepts of probability and statistics. In the latter, the best decision depends entirely on the probabilities we assign to various events. In this sense, the three examples illustrate a critical feature of most decision problems: we have to make decisions on the basis of what will probably happen in the future.

Election-night predictions

In the United States, we are now accustomed to relatively brief election evenings. Shortly after the polls close in any state, the television networks are predicting the winner and estimating with astonishing accuracy his final percentage of the vote. A major problem arises because of the speed with which a winner is declared; in a national election, the winner is usually known before the polls have even closed in the western states.

This system of high-speed prediction is quite new. As late as the 1964 elections, each network collected vote information independently. Then the News Election Service (NES) was organized; now the three networks and two wire services report for each state only the vote totals which come from NES. (For this reason, the totals shown to the public on the networks are the same during the evening). Thus, there is unified coverage of the 175,000 precincts and 3000 counties throughout the country.

Freed of the need to collect vote totals, each network focusses its efforts on collecting data appropriate for prediction of winners. The prediction is based for each state on 50 to 150 key precincts. For each, a mathematical model takes into account the voting pattern in previous elections, the pre-election polls, and the returns on election night as phoned in by network observers.

From the results from these key precincts and a study of the state as a whole, the computer develops a prediction. This model also takes into account the effects of the weather and other factors influencing the voter turnout. The computer prediction then represents an optimum decision based upon the probability that the behavior of voters in other areas will follow those in the key precincts. Actually a winner is not predicted early unless the model shows at least 3 or 4 percent difference in order to allow for errors.

Thus, the only circumstance under which the computer-model prediction will err would be if there were significant, unanticipated factors which invalidate the choice of key precincts. For instance, if the population of the state is 20% Catholic, one candidate is heavily favored by Catholic voters, and the key precincts contain very little Catholic population, the computer prediction might be grossly in error. Hence, the major effort before the election is to ensure that the key precincts chosen are really representative.

Route for auto travel

The second example involves the selection of a route for auto travel from a finite set of possibilities. A specific example is shown in Fig. 5-5, which is a map of the main parkways from Huntington, Long Island to LaGuardia Airport. The driver, situated at the start point, has a number of decisions he must make to determine his route to LaGuardia Airport

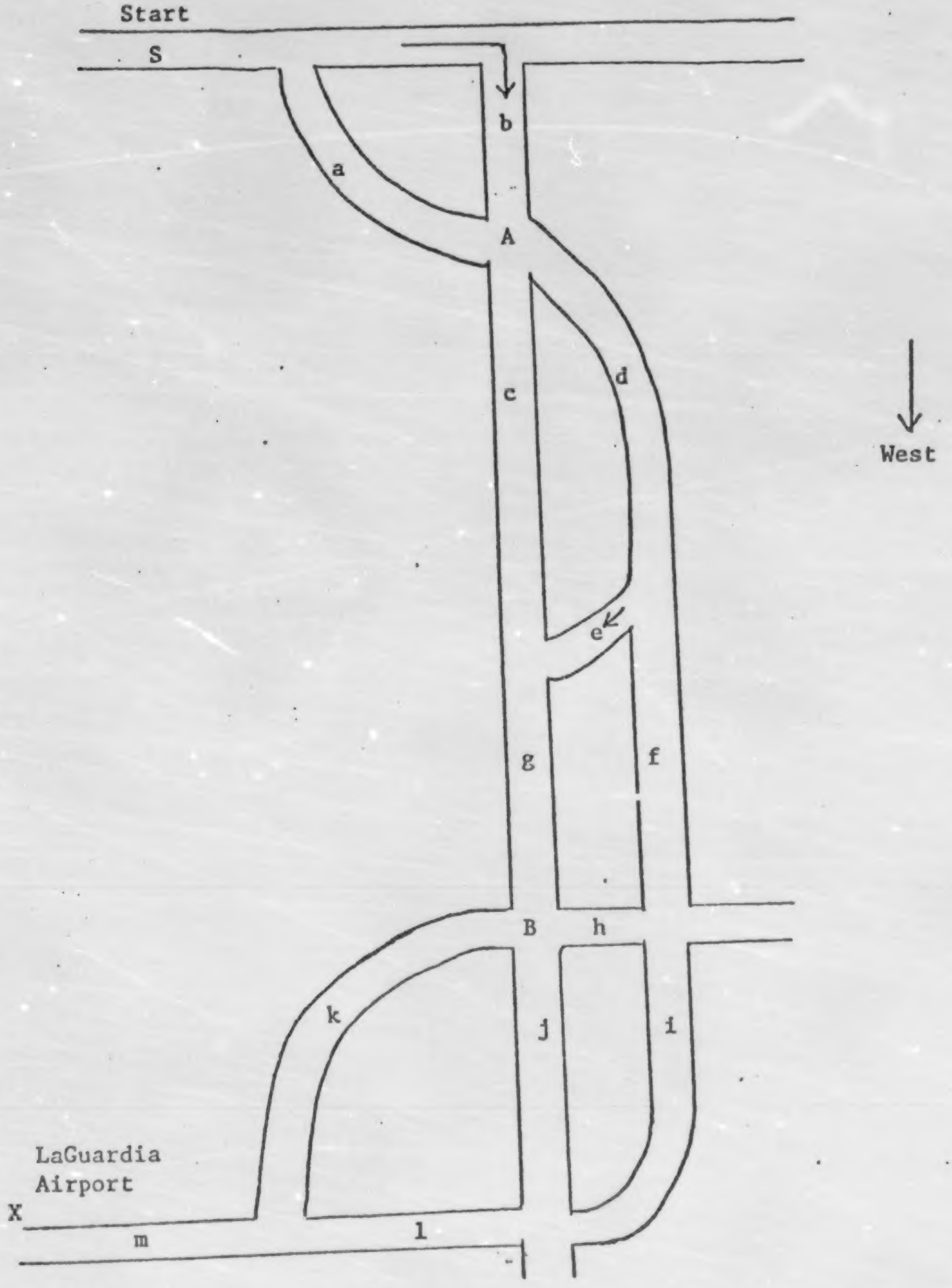

Start
S
b
a
A
c
d
West
e
g
f
B
h
k
j
i
LaGuardia
Airport
X
m
l

Fig. 5-5 Major parkways from Huntington, N.Y. (start) to LaGuardia Airport

(35 miles distant).

There are two principal routes west: segments bcgjlm and adfilm. At several points, however, cross-overs are possible. As a result, the relatively simple decision problem actually has 18 different routes which we can ennumerate.

acgjlm
acgkm
acghilm
adegkm
adegjlm
adeghilm
adfhkm
adfhjlm
adfilm

bcgjlm
bcgkm
bcghilm
bdegkm
bdegjlm
bdeghilm
bdfhkm
bdfhjlm
bdfilm

The complexity of the problem is noteworthy. Even though we have omitted any possibility of travel on local streets, there are 18 routes from which the driver must choose. (Actually, even if we consider only parkways, the Long Island network is much more complex than shown here; we have simplified by considering only routes which are very close to the minimum mileage).

In any decision process of this type (the choice of one from a number of alternatives), we first search for a way to simplify the problem. Is there any way to break the complex problem down to a set of simpler problems? Two decisions, one with 2 alternatives and one with 9, are easier than one decision with 18 alternatives.

Fortunately, this problem of Fig. 5-5 can be immediately decomposed into two successive decisions:

How to travel from S to A? (2 alternatives)

How to travel from A to the airport? (9 alternatives)

We have viewed the problem as a *multi-stage decision* in the language of decision theory. In other words, we view the travel in two stages, rather than as a single, overall trip. First the driver chooses between *a* and *b*; then he chooses his route from A to the airport.

This decomposition or breaking down of a problem can often be carried much farther than in the above example. For instance, if the driver knows that segment *i* is almost always seriously congested, he may decide at the outset to reject any thought of travel on this portion. Then the problem is shown in Fig. 5-6.

Now there are two intermediate points, A and B, through which he must pass. Hence, the problem can be broken down into a *three*-stage decision:

S to A	2 alternatives	a b
A to B	3 alternatives	cg deg dfh
B to Airport	2 alternatives	km jlm

The driver makes three decisions, at S and A and B. The three small problems are much simpler than one large problem with 12 alternatives (if i is omitted in Fig. 5-5, there are 12 rather than 18 routes, as indicated by the listing at the beginning of this subsection).

Once the problem is structured in this way, how does the driver actually make a decision? For an intelligent decision, he should have information about the congestion existing along each of the alternative routes. For instance, as he leaves from S, he has to decide between *a*

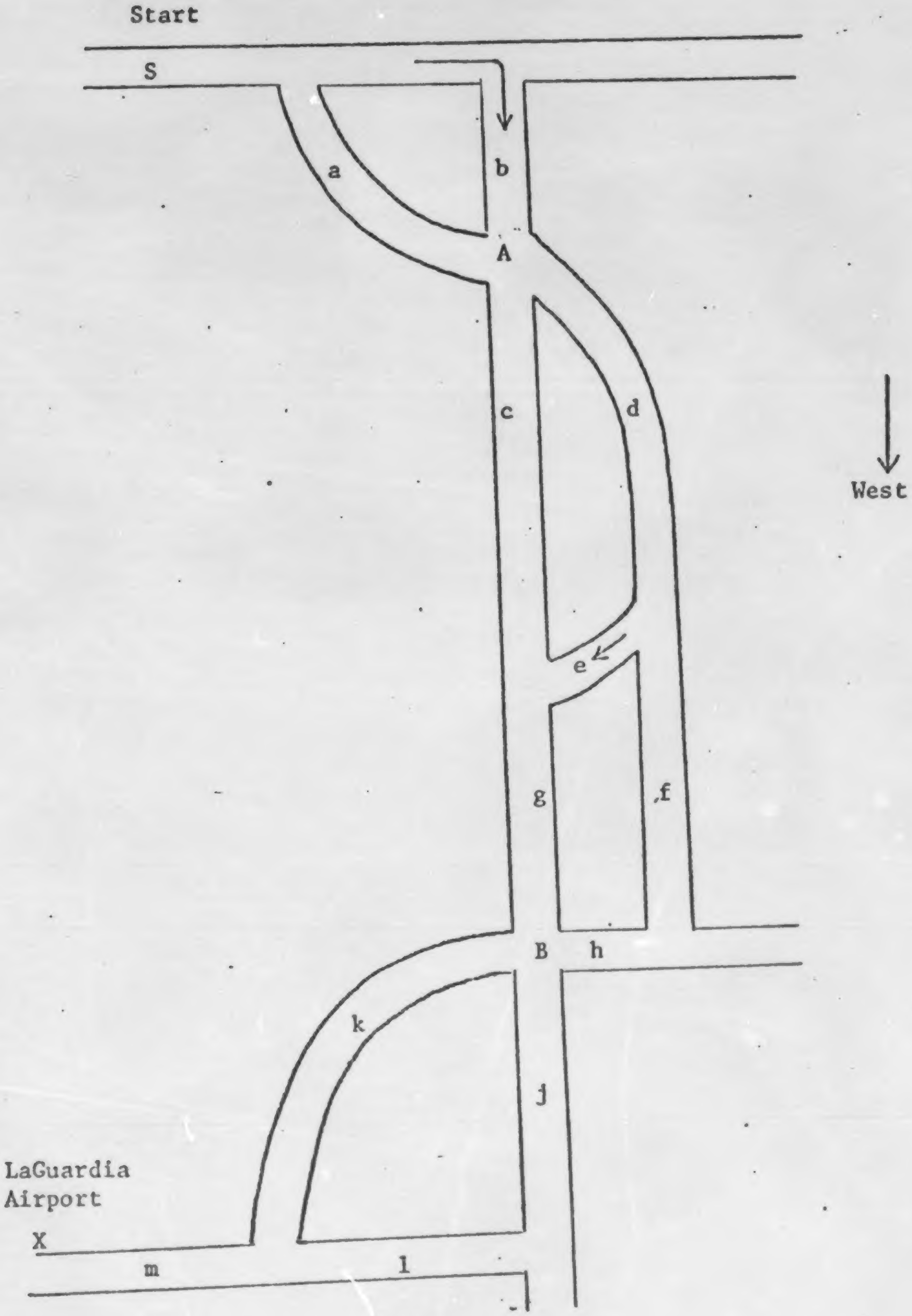

Start
S
a
b
A
c
d
West
e
g
f
B
h
k
j
LaGuardia
Airport
X
m
l

Fig. 5-6 Possible routes with segment i in Fig. 5-5 omitted

and b. He needs to know the traffic congestion on each of these segments. In most such situations, the only information the driver has is that derived from his past experience in driving these roads at this time of day and under these weather conditions. Unfortunately, this experience may simply tell him that there are tremendous variations from day to day.

At several locations around the United States, there are now appearing attempts to give the motorist more information on which to base a decision. When an urban parkway may cost as much as $30 million per mile, quite elaborate communication systems can be added at a cost which is relatively trivial. Thus, we are beginning to find electronic systems which measure traffic flow rates further down the highway. Then at the decision point S, the motorist is informed by an electric sign that traffic is moving at 40 mph on a, at 15 mph on b. The signs can also indicate the location of accidents, disabled cars, flooding, fog, or other factors affecting traffic flow.

There is one serious problem with such information aids to decision making. Unless the signals sent to the motorist anticipate or predict motorist reaction, the system is apt to be unstable. If the travel time on route a is much less than b, most motorists will select a. Within a few minutes, a will be terribly congested while speeds on b rise rapidly.

The system can be designed to anticipate this human response and, for example, warn against travel on a just before congestion there reaches a serious level. Alternatively, some human drivers will learn to behave in just the opposite of the normal response and will thereby benefit maximally from the system.[1]

The dominant characteristic of this route-selection decision problem, however, is that it is a probability problem. Even after the problem is broken down into a set of simple problems, each of these must be viewed in terms of probability. Congestion and accidents or other traffic tie-ups can not be predicted. Hence, as this driver considers his choices, he must decide (from his experience plus any information available) which route is probably favorable.

Just because the decision is probabilistic, he can do no more than optimize his probability of reaching the airport in the minimum time or with the least annoyance. Regardless of how intelligent the decision, there is always the possibility of a disastrous delay on any route he may choose. If he drives the route frequently, he can only hope to make decisions which over the long run will result in the most satisfactory trip.

5. One Man, One Vote

During the last decade, widespread publicity has been given to the "one-man, one-vote" concept -- the idea that in an election or voting situation, each vote should carry equal weight. The problem was first highlighted in the legislatures of the more populous states which had both rural and urban areas. Because of the historical emphasis on property ownership as a criterion for voting elegibility and the strongly agricultural character of the country as recently as the 19th century, the rural parts of the state often dominated the legislature. It was not unusual to find a rural legislator representing 5,000 people, while his city colleague spoke for 40,000. While the total state population might be more urban than rural, legislative actions strongly favored the rural sections and often resulted in only passing attention to the increasingly critical, city problems.

---

[1] The desirability of responding irrationally to certain information appears often in system studies. One of the best examples is in career selection. The best time for an entering college freshman to choose a certain career is often when there is maximum publicity about unemployment and oversupply of personnel. Four years later when he graduates, there is likely to be a severe manpower shortage with resulting multiple job opportunities and high salaries. Unfortunately, career counselors often seem to overlook the importance of the oscillating nature of employment and the time delay inherent in education.

A series of court decisions led to the one-man, one-vote concept. The legislature had to be apportioned in such a way that each representative speaks for approximately the same number of people. This same idea then was extended to a wide variety of decision processes, with general acceptance of the principle that decisions made by the public or its representatives should be based on the one-man, one-vote concept as far as possible.

Very often in decisions relating to socio-technological problems, the concept of one-man, one-vote does not result in a fair treatment of all interested parties. The decision process is strongly affected by bloc voting, when a sub-group has a common interest and always votes together. In order to have equal representation of each voter, we have to anticipate the way blocs will form and set the decision-making rules accordingly.

In the following paragraphs, we consider one example which illustrates the problem. Our goal here, however, is not the study of the effects of bloc voting, but rather the recognition that decision-making in competitive situations has to be organized very carefully if we are to be sure that the ultimate decision does represent the best possible contributions from all interested parties.

Potomac River ecology project

The Potomac River basin is of primary interest to residents of

Maryland Pennsylvania Virginia
West Virginia District of Columbia

Recognizing the inter-state problems and the importance of ecological planning, Congress passes legislation establishing a commission to develop policy and regulations to control the use of the water, the allowable effluents from towns and industries along the river, and the development of parks and recreational areas.

The framers of the legislation believe that a commission of 11 members is the largest workable size. One member represents the U.S. government and is appointed by the President. The other ten members are elected by the people directly concerned according to population, with the result that the membership is allocated by states:

Maryland - 3 Pennsylvania - 1 Virginia - 3
West Virginia - 1 District of Columbia - 2

The authors of the bill setting up the commission want to be sure that no one group dominates the decisions, so the requirement is imposed that any decision must be made by a 2/3 majority of the full group, or by 8 votes.

Thus, the commission is established with careful regard for the one-man, one-vote concept, and the rules are set to be sure that any decision does represent the studied opinion of at least 2/3 of the members. The legislation seems to be ideal.

The problem of blocs

Shortly after the commission starts operating, the two representatives of the District of Columbia and the Presidential appointee are dismayed that they have no real voice in most of the decisions. What has happened is that the interests of Maryland and Pennsylvania are usually the same, and these delegates regularly vote together. Similarly, Virginia and West Virginia attitudes are parallel. As a consequence, there are in actuality four voting blocs:

Maryland - Pennsylvania	4
Virginia - West Virginia	4
District of Columbia	2
U.S. representative	1

With 8 votes needed for action, a decision is only possible if there is agreement by the first two blocs. How the D.C. and U.S. delegates vote is totally immaterial!

Frustrated by their helplessness, the D.C. and U.S. representatives ask Congress to amend the legislation so only a majority vote is needed to pass a measure. The other 8 delegates object, claiming that this would give unjust power to the minority members. Is this true?

Analysis

To decide on the fairness of the proposed rule that a majority is needed to reach a decision, we might look at all the ways a proposal can pass. Those voting "Yes" can be :

11-0

Md-Pa (4)
Va-WVa (4)
DC (2)
US (1)

10-1

Md-Pa (4)
Va-WVa (4)
DC (2)

9-2*

Md-Pa (4)
Va-WVa (4)
US (1)

8-3*

Md-Pa (4)
Va-WVa (4)

7-4*

Md-Pa (4)
DC (2)
US (1)

7-4

Va-WVa (4)
DC (2)
US (1)

6-5*

Md-Pa (4)
DC (2)

6-5

Va-WVa (4)
DC (2)

There are 8 ways a measure can pass under the majority rule.

In how many of these does the Md-Pa block have the deciding vote? That is, in how many combinations does the "Yes" vote of Md-Pa result in passage? There are four cases shown above by the astericks. (In the 11-0 situation, reversal of the Md-Pa vote would not change the outcome). In other words, the Md-Pa bloc is the decisive vote in four situations.

Similar calculations reveal that each bloc is critical in the following number of cases:

Md-Pa	4	DC	4
Va-WVa	4	US	0

The vote of the U.S. delegate is now immaterial; he is effectively disenfranchised. In addition, the two delegates from the District of Columbia have as much importance as either of the four-vote blocs. Thus, the majority rule does give too much power to the minority members, at least the two from D.C.

This example suggests some of the perils in a third-party movement in this country. If in the Senate the two major parties are about even in strength, a few Senators from the third party can have enormous power. The importance of the third party becomes even greater if votes are traded -- the third party agreeing to favor a measure in exchange for a pledge of support by one majority party on another bill.

The analysis above, showing equal power for the three blocs, is based on the assumptions that each of the eight possible voting patterns is equally likely and that complete bloc discipline is realized (each bloc always votes unanimously). In the actual operation of the commission, there certainly are some issues on which blocs split, but the analysis does describe all questions for which the blocs hold.

An alternative proposal

Unhappy with the disenfranchisement of the U.S. representative, the Congress instead amends the law to require 7 "Yes" votes for a decision. Is this a logical and fair basis? Similar analysis shows that each bloc now has the "swing" vote in the following number of cases (we simply omit the 6-5 votes in the earlier listing):

Md-Pa	4	DC	2
Va-WVa	4	US	2

The ideal pattern would be 4-4-2-1, proportional to the number of delegates. The principal discrepancy is in the added power of the U.S. representative, but presumably he is a disinterested member concerned with the public interest.

Final comment

This example[1] is included primarily to emphasize the fact that the rules under which decisions are made (here the votes required for passage) may determine the extent to which the decisions are responsive to various interests and attitudes. Energetic, vocal, and persistent groups have always had considerable success, particularly when the majority is apathetic. Rules for decision-making can completely distort the logic.

---

[1] The study is known as the Shapley analysis, named for the social scientist who first presented this approach in the 1950's.

6. Competitive Decision-Making

The problem of the Potomac River basin is an example of competitive decision-making. There we looked at competition among four separate parties or voting blocs, with each anxious for its own influence in the decisions. There is also a class of decision problems in which we focus on competition between two parties, each striving to benefit at the cost of the other.

Such face-to-face competition arises when there are two principal suppliers of a single product or service, and they are competing for a fixed market. For example, if only Delta and Eastern Airlines fly from Baltimore to Atlanta and the annual number of passengers is approximately fixed by business and pleasure needs of the public, the two companies are in direct competition. What one gains in customers, the other loses. In developing a strategy to try to increase its share of the market, Delta must decide on an advertising policy, how to set fares within governmental regulations, when to schedule flights, and how much to invest in improved service. Each of these four directions is also being pursued by the competitor, Eastern.

What Delta decides to do is determined (within cost constraints) by its estimate of the actions Eastern will take. Delta tries to find that strategy which maximizes its expected gain under the assumption that Eastern, with equal intelligence, will choose its strategy to minimize Delta's gain.

We have here the basic elements of a two-person, zero-sum game: two-person because there are two competitors, and zero-sum because what one loses, the other gains. The solution of problems of this sort comes under the mathematical subject of game theory, first discussed in detail in the book "Theory of Games and Economic Behavior," by John von Neumann and Oskar Morganstern.[1] When a finite, quantitative model can be determined, game theory problems can be solved. This model lists the specific, alternative strategies open to each competitor -- for example, one strategy for Delta might be to spend $2 million more on advertising, schedule flights at 9 a.m., and make no change in fares or service. The model must also include the payoff to Delta which results from each possible combination of Delta and Eastern strategies.

A zero-sum game

The elements of a game are most easily illustrated by a simpler example, in which each competitor has only three possible moves. For example, Barbara and Tom are playing the following game. Each secretly writes down one of the three letters (A, B, or C) on his piece of paper. Then they both reveal their letters. Figure 5-7 shows the amount Tom then pays Barbara, according to the combination. If Barbara has written B and

---

[1] The third edition of the book was published by Princeton University Press in 1953. J. von Neumann was the brilliant mathematician who also conceived the idea of the stored program for a digital computer, the concept which allows a computer to be used for many different problems.

		Tom		
		A	B	C
Barbara	A	0	45	-15
	B	15	0	15
	C	-15	45	0

Fig. 5-7 Payoff matrix -- payoff to Barbara for each of nine possible combinations

Tom A, we are at the first entry of the second row. The payoff to Barbara is 15: Tom gives Barbara 15 cents.

While this may not seem like a very exciting game, it is a simple analog to the Delta-Eastern competition. In that realistic conflict, each airline has dozens of possible moves or strategies from which to choose, and the payoffs are very difficult to calculate (one must estimate the change in passengers after each possible combination of strategies).

For the trivial game modelled in Fig. 5-7, let us look at Barbara's reasoning as she tries to decide what letter to write. If she assumes Tom plays completely at random (he may just arbitrarily choose A, B, or C, with equal probability), and if she plays A, she can expect to win 0 one third of the time, win 45 one third of the time, and lose 15 one third of the time. Hence, if she plays A, she will win on the average 10 cents. In the same way, if she plays B, she can expect to win 10 cents. If C is her choice, again 10 cents is her average winning. It seems it makes no difference what she chooses.

She knows Tom is intelligent. If he thinks she will play A, B, and C at random, he will always play A or C. If he plays either and she plays at random, Barbara's average winning will be zero. Suppose he does play only A or C; then obviously she should play only B, so she will average 15 cents every time the game is played.

But Tom reasons as follows: Barbara will guess I'll play A and C; she'll play B; then I should play B to drop her winning to zero. If she knows I'm going to play B, she'll switch to A or C to win 45 cents.

As you can see, we do not seem to be moving toward a solution. We are in the endless round robin: he thinks that she'll think that he'll think, and so forth.

The mathematics of game theory provides the best solution for each player, as well as the average value of the game to Barbara. This solution assumes both players are intelligent and both understand game theory. There is no necessity for us here to go into the mathematics by which the best way to play can be determined;[1] we are instead interested here in the type of decision problems which can be solved with game theory and the nature of the solution.

In the present problem described by Fig. 5-7, the best way for Barbara or Tom to play is as follows:

$$\text{Barbara should play} \quad \left.\begin{array}{l}\text{A one ninth}\\ \text{B seven ninths}\\ \text{C one ninth}\end{array}\right\} \text{ of the time}$$

---

[1] An explanation written for the layman is included in the book "The Compleat Strategyst," by J. D. Williams, McGraw-Hill Book Company, New York, 1965. The author includes an amusing series of examples which show some of the unexpected results of game theory.

Tom should play A, B, and C one third of the time each.

The expected value to Barbara of each play is 10 cents.

Each competitor should pick A, B, or C by a random, unpredictable process, but according to the above probabilities.

Before each play, Barbara should make her selection to meet the above probabilities (A 1/9, B 7/9, C 1/9), but with the choice made at random. If she has a table of random numbers available, she might decide to use a row of single digits. Before looking at the table, she would decide the ten digits (which are equally probable) will be assigned as follows:

1 will correspond to A

2-8 inclusive will correspond to B

9 will correspond to C

0 will be disregarded

She is using only nine digits, so that each one has a probability of 1/9. The desired probabilities of 1/9, 7/9, and 1/9 are now obtained by choosing one digit, seven digits, and one digit, respectively.

Barbara now opens the book of random numbers and picks one row of digits:

4 7 2 0 2 9 3 3 1

The first eight times she plays, she selects

B B B   B C B B A

in that sequence. If the game continues, she uses different sequences from the table of random numbers. Tom plays in the same random way, although his probabilities are 1/3, 1/3, and 1/3.

When the game is played in this way, neither competitor can predict what his opponent will do on any given move. If Tom would begin to follow a pattern and Barbara could predict with modest success when he would play B, she would be certain to choose A or C in the hope of winning 45 on that play.

Non-zero-sum game

Many of the game or competitive situations which occur in the real world are not zero-sum. What one side gains is not necessarily the same as the other side's loss. Even in the Delta-Eastern competition for Baltimore-Atlanta passengers, advertising by either airline is partly designed to increase the total market. Hopefully, more people will be convinced to fly, with benefits for both of the competitors. When the game is not zero-sum, the standard solution techniques of game theory fail. While we can sometimes find a solution, the determination of the best strategy for each player often depends on searching through the possible solutions.

The non-zero sum game is illustrated by one of the classic problems of operations research. Two men each have an ice cream wagon which they bring each morning to a mile-long beach. During the day, customers or sun-bathers are distributed uniformly along the beach. When a vacationer wants ice cream, he goes to the nearest vendor, since there is no real difference between the two.

When the first vendor arrives in the morning, where should he locate? Where should the second man then settle with his wagon? We assume that there is no cooperation between the two: each man doesn't trust the other, and each acts solely to maximize his own sales.

A little thought shows that the two vendors should locate next to one another exactly in the middle of the beach (Fig. 5-8). If A, the

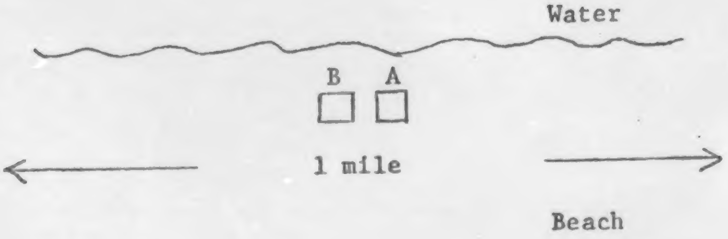

Water
B
A
1 mile
Beach

Fig. 5-8 Location of ice cream wagons A and B in a strictly competitive situation

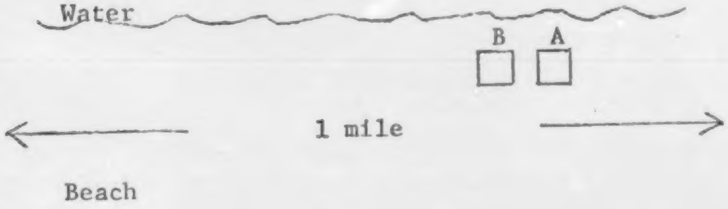

Water
B
A
1 mile
Beach

Fig. 5-9 Final location if A selects any point other than the center

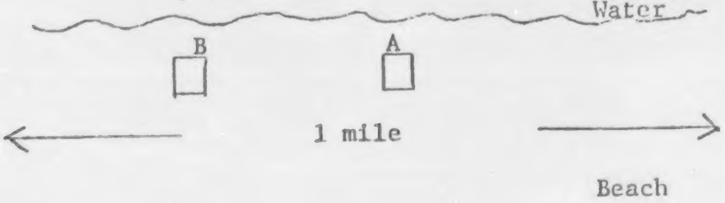

Water
B
A
1 mile
Beach

Fig. 5-10 Final location if A chooses the middle and B makes a non-optimum choice

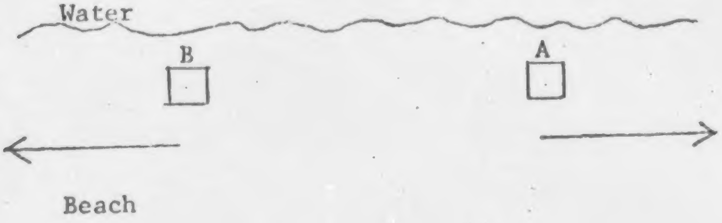

Water
B
A
Beach

Fig. 5-11 Solution best for the customers and equal to the optimum for both A and B

first to arrive, makes any other choice, B will then settle just on the side of A toward the longer expanse of beach (Fig. 5-9). If A picks the middle and B locates anywhere but beside him, A will be serving more than half the customers (Fig. 5-10).

When the two sides are strict competitors, the game or problem leads to a ridiculous solution, and certainly one which does not give the best possible service to the customers. From the standpoint of the customers, the optimum solution is shown in Fig. 5-11. Each vendor is in the middle of "his" half of the beach. The customers have shorter distances for ice cream on the average, yet each vendor serves half the people. In actuality, Fig. 5-11 is likely to be better for both vendors than the competitive solution of Fig. 5-8, since the mere fact an ice cream wagon is nearby for more people should increase sales.

Why is the solution of Fig. 5-11 so difficult to achieve in a competitive situation? It requires honest cooperation between the two competing forces -- a cooperation which has to be based on mutual trust. When A arrives in the morning, he takes his possibly bad position in confidence that B will not try to take advantage of his trust. During the day, both resist the temptation to move gradually toward the center of the beach. This is a non-zero-sum situation: both benefit by disdaining what looks like an advantageous possibility.

This rather trivial example has many counterparts in the real world. One of the most obvious is the fierce competition among three airlines (TWA, American, and United) for passengers on the lucrative New York-Los Angeles trip. In an attempt to attract a large fraction of the customers by having flights at convenient times, each airline runs more flights than necessary (on the average, less than half the seats are filled) and the three schedules tend to be startlingly similar--with three flights taking off almost concurrently, then a gap of several hours before the next three.

Part of the problem here has been the federal anti-trust laws and the concept that the public is best protected by complete competition among firms. It is only in the past year that the government has allowed the three airlines to talk together in an attempt to agree on schedules which would improve both service and profits.

In a broader sense, non-zero-sum decisions dominate our society, as G. Hardin has pointed out in his classic article.[1] An extreme example illustrates the point. Each author pays $10 per month for rubbish collection from his home. From a selfish standpoint, we would benefit if we each simply threw the day's rubbish out the car window while driving to work each morning (and while making sure we were unobserved). The damage to the environment would be negligible, we would save $20 monthly, and the quality of life in our area would not deteriorate in any real way if just the two of us followed this practice.

The problem is obvious. If everyone, or even a significant fraction of the populace, acts in this same selfish way, we all would shortly be surrounded by rubbish. In order to find a social optimum -- a decision or strategy which results in the best total picture -- we each must make decisions which are far from optimum.

---

[1] The Tragedy of the Commons, Science, Dec. 13, 1968, pp. 1243-8

As we consider different socio-technological problems, we find again and again this non-zero-sum character. The total system only works if the individual is willing or can be forced to relegate his personal benefit to a secondary position compared to the total, social welfare. Just as in the ice-cream-vendor problem, success depends on individual trust and confidence in others. Whether the concern is obeying traffic regulations or conserving energy, we must design total systems in which the individual is willing to accept decisions which are obviously not the best for him personally.

7. Decision Problems Without Solutions

In this chapter, we have attempted to look at some of the principal characteristics of the decisions which accompany socio-technological problems. There is no attempt to cover completely the topic of decision-making,[1] and much of what engineers call decision theory[2] is omitted entirely.

---

[1] Several other topics are discussed in Chapters 2 and 3 of "Man and His Technology," McGraw-Hill Book Company, New York, 1973. There are also many articles in the public press which recount specific accomplishments and current decision activities. For example, one of the more elaborate decision problems is described in Uri Shamir, Optimizing the Operation of Israel's Water System, Technology Review, June 1972, pp. 41-48.

[2] Decision theory is concerned with the problem of finding the best interpretation from data which is confused by noise and uncertainty. When pressure within the eye is measured, for example, in an attempt to detect the disease of glaucoma, what conclusions can we draw on the basis of the knowledge amassed in the past about similar measurements on both healthy people and glaucoma victims?

Before leaving the topic to return to specific problem areas, we should cite one additional example which points out the possibility that a decision problem may have no solution at all, even though the problem is taken from the real world. This idea of no solution is strange to most scientists and engineers, who are used to working with well defined physical systems in which problems have a single correct answer.

The city mayor has a small amount of money available to try a new program to improve life in the city. Three different programs are proposed:

T -- traffic-flow improvement
A -- air-pollution reduction
R -- rubbish-collection modernization.

Studies convince him that the money can not simply be divided among the three alternatives. If there is to be any progress with the small number of dollars available, they must be put into a single program. Which one?

Fortunately, the mayor has a staff of system analysts. He asks them to determine the probable effects of each of the three options. After detailed analysis, they report that the city is such a complex system, they just can not predict with confidence what the effect of any one program will be. There are too many unknown influences at work.

The most they can say is that, if program T (traffic) is adopted, any of three things may happen, with equal probabilities. There may be much better traffic flow, with an improvement in the quality of city living by an amount they estimate to be +40. There may be no change (value 0). Finally things may actually become worse as the smoother traffic flow draws more cars into the city, air pollution rises, parking is hopeless, and so on; to this last possibility they assign the value -10.

For each of the other suggested programs, system analysts foresee three, equally probable outcomes. The values assigned to each are summarized in tabular form for the mayor:

T	A	R
+40	+30	+15
0	+20	+10
-10	-20	+5

The mayor is himself an amateur system engineer. He thanks the advisors and heads home to make his decision. First, he compares T to A. Since T has three possible outcomes and A does likewise, there are nine equally probable combinations:

T	$\boxed{40}$	$\boxed{40}$	$\boxed{40}$	0	0	$\boxed{0}$	-10	-10	$\boxed{-10}$
A	$\boxed{30}$	$\boxed{20}$	$\boxed{-20}$	30	20	$\boxed{-20}$	30	20	$\boxed{-20}$

In the five boxed cases, T is better than A. Aha, the mayor concludes, 5/9 of the time T is better than A. Obviously I would choose T in preference to A. Now I've simplified the problem to a choice between T and R.

Similar comparison of T and R gives

R	15	$\boxed{15}$	$\boxed{15}$	10	$\boxed{10}$	$\boxed{10}$	5	$\boxed{5}$	$\boxed{5}$
T	40	$\boxed{0}$	$\boxed{-10}$	40	$\boxed{0}$	$\boxed{-10}$	40	$\boxed{0}$	$\boxed{-10}$

In the six boxed cases, R is better than T. Thus, there is a 2/3 probability that R is preferred over T. Clearly, the mayor should choose R. The decision problem is solved.

Chortling over his success in optimization and decision techniques, the mayor makes one mistake: he decides to compare R and A:

A	30	30	30	20	20	20	-20	-20	-20
R	15	10	5	15	10	5	15	10	5

To his dismay, there is a 2/3 probability that A is better than R.

This particular decision problem has no solution:

T better than A better than R better than T

No matter what the mayor chooses, his opponents can always find a program which would have been better.[1]

The important feature of this problem is the total lack of a solution. The first reaction of a system analyst is often to attack the validity of the problem on the basis that the three sets of values assigned to T, A, and R are not defined. What these values mean is really immaterial, however. If values can be determined in any sort of acceptable and logical way, they might be the numbers we have used here as an illustration.

In other words, this example illustrates that there are decision problems for which a solution does not exist. This chapter opened with a discussion of the drug control problem, in which so much of the quantitative model is unknown that no logical decision is possible. We close with a probability decision problem in which the model is known completely, yet again there is no solution.

These two examples represent extreme cases; in the great majority of decision problems which arise, we can find a solution, even though our confidence in the value of the solution depends on the quality of the model.

---

[1] This type of behavior, called the intransitive property, was first discussed in decision problems in the 1950's by Prof. K. Arrow of Harvard, an outstanding economist who received the Nobel Prize in Economics in 1972.

CHAPTER 6

THE ENERGY CRISIS

Centuries ago, people living in the tropics learned that wet cloth hung over the windward side cooled the air entering the house. Air conditioning became a science with the work of Willis H. Carrier of Buffalo, N.Y. in the first decade of the present century. Although the earliest public use was in a movie theater in 1922, air conditioning of office buildings, stores, and homes did not reach a large scale until the 1950's.

By the early 1970's, air conditioning is a significant contributor to the energy crisis besetting this country. By 4 p.m. on a hot summer day in New York, the electric utility (Con Edison) has very likely reduced the voltage of the electric energy delivered to the city in order to lower power demands -- a step called a "brown-out" in contrast to a "black-out" when power is lost entirely. The 4 p.m. hour is particularly critical since air conditioners are operating at maximum capacity, subways and trains are beginning the rush hour, and a start is often made on cooking the evening meal. By five o'clock, offices are closing, the outside temperature is falling, and the total power demand begins to drop.

As the demand for electric energy continues to double every nine years, there is not a corresponding growth in capacity. In the last few years, environmental interests and the public have opposed new generating stations with increasing success. Since a half dozen years are required to plan, build, and bring into operation a new plant, the energy shortage promises to be critical for some time into the future regardless of any new construction authorized now.

In the following sections, we want to look in more detail at this energy crisis. How serious is the crisis? What are the alternatives open to the public? How can intelligent planning be done to ensure that a satisfactory compromise is reached between the energy desires of the public and the costs to the environment? What are the possibilities for realizing President Nixon's 1971 goal of developing new, clean sources of energy?

1. What Is Energy?

In science texts, energy is defined as the capacity to do work. When a 160-pound man climbs a flight of stairs, he gains potential energy by virtue of his height; the gravity force pulling him downward means that work could be done as he falls. Similarly, a mass in motion possesses kinetic energy; work can be done as the mass is brought to a stop.

About 1800, scientists discovered that heat is one form of energy. Humphry Davy showed that two ice cubes can be melted by rubbing them together even when the temperature of the environment is below freezing. Part of the work done in the rubbing motion is converted to the heat form of energy. Actually it was 50 years later before James Joule measured accurately the quantity of heat corresponding to a given amount of mechanical energy.

There are many other types of energy. Electrical energy is used to drive most of the home appliances; chemical energy from gasoline drives an automobile; sound energy carries information from man to man. The human body converts the chemical energy in food into a variety of forms: heat to maintain the body temperature, mechanical energy which allows the heart to circulate the blood and the man to exercise or move, and electrical energy to carry signals through the nervous system.

While at first glance the idea of energy may seem simple, it is actually one of the most difficult concepts in science. One of the problems is that energy is difficult to define precisely. If an airplane is moving at 400 mph, it clearly possesses kinetic energy: work can be done as the motion stops. If we are riding at the same velocity alongside the airplane, however, it seems to us to be motionless, hence to have no kinetic energy.

Often when we try to use the energy concept for analysis to understand some physical phenomenon, we find that it is extremely difficult to determine how the energy is converted from one form to another. For example, there has been a long-term interest in the energy required for various animals to move by running, flying, or swimming. In spite of the studies, there are still many unanswered questions.

When a man runs, the energy used up can be measured by the change in his metabolic rate, indicated by the net oxygen intake. Since the energy needed to overcome air resistance and friction along the ground is negligible,[1] essentially all the energy is used in the motion of the joints and muscles and in accelerating and decelerating the parts of the body. The determination of how the energy divides among these various functions is an exceedingly difficult problem, with the result that the running process is really not understood at the present time. Obviously more energy is required to run uphill than on the level. But it is still unclear whether running downhill allows the muscles to use some of the potential energy which is lost.

---

[1] At top sprint speeds, air resistance does have a slight significance, so it is appropriate to disallow sprint records under strong, favoring wind conditions.

If various forms of locomotion are compared for different animals,[1] running requires much more energy than flying even though flying requires considerable effort just to stay aloft. This comparison reveals why a migrating bird can fly non-stop for more than 500 miles, while a ground animal of similar size certainly can not travel the same distance without pausing to eat and drink. In turn, swimming requires much less energy than flying, when we compare animals for which these are natural methods of movement. When man is in the water, travel requires far more energy than running because of the enormous waste of muscular and joint energy as the human being tries to adapt to this unnatural way of locomotion.

The concept of energy appears throughout science and engineering. When we speak of the energy crisis, the central theme of this chapter, we are focussing on the energy used to drive machinery in production, heat and air condition, and transport people and materials. We are interested in the energy which man extracts from natural resources (coal, natural gas, and petroleum) or nature's sources (waterfalls, tides, winds, and sunlight), and which he then converts to forms he can control for his desired purposes. We focus on a relatively small part of the total energy concept, and in particular we omit entirely the physiological uses of energy by man himself.

---

1. Knut Schmidt-Nielsen, Locomotion: Energy Cost of Swimming, Flying, and Running, Science, July 21, 1972, pp. 222-228

2. Problems of the Energy Crisis[1]

When the public media discuss the energy crisis, they usually focus on two principal problems associated with the rapid growth in energy use:

Depletion of natural resources

Environmental effects

In one sense, neither factor is really a problem.

When the authors were in college around 1940, there was great concern expressed about the future disappearance of coal and oil. Dire predictions were made that by 1970 the United States would be in a catastrophic position. Today we hear similar concern expressed for as early as 1990. In parallel with this concern for resources, there is widespread recognition that many of the environmental problems, and especially air pollution, are closely related to the increasing generation of electric energy and the growth in transportation (both automobile and aircraft).

Both of these problems are solvable. The world, and even the United States, has sufficient natural resources (especially coal) to last far enough into the future to allow the technology to advance to the point where solar energy can be used. We now have the technology which allows the control of air and thermal pollution and other environmental effects. Then why is there an energy crisis?

---

[1] The energy crisis has been discussed extensively in journals and newspapers as well as on television. Perhaps the best references from the viewpoint of the engineer or scientist are the series of articles entitled "Energy Technology to the Year 2000," appearing in the Technology Review October/November 1971, December 1971, and January 1972. In September of 1971, Scientific American had a special issue on Energy and Power. A large percentage of the articles in Environment magazine are devoted to the subject, particularly the environmental effects. Finally, there are many governmental studies; one interesting report analyses in detail the ways in which energy is consumed: "Patterns of Energy Consumption in the United States," by the Stanford Research Institute, GPO Stock Number 4106-0034.

The crisis arises because the U.S. social system has developed on the basis of a particular pattern of energy use. Natural gas is used widely in heating as well as industrial processes; almost all vehicles are powered by gasoline; the energy needs and expectations of society are fixed. Of equal importance, the costs of different forms of energy have been established over the years and have determined not only the form of energy use, but such socially critical factors as the location of industrial concerns.

As natural resources are depleted, the oil and gas companies (for example) have to turn to less accessible and more expensive sources. As the public demands stricter environmental controls, the generation and distribution of energy in usable forms become more costly. Any significant change in the cost of one type of energy or energy in one location results in enormous disruption of the social system. The threat of such a disruption is the cause of the energy crisis.

The two reasons usually listed for the crisis,

Resource depletion Environmental effects

are important, but because these factors mean that there is a threat of disruption of the existing pattern of energy use.

The consequent social effects can be severe. For instance, many of the larger cities are finding an out-flow of manufacturing companies which represent jobs for the semi-skilled members of the working population. The decision by a concern of whether to move out of the city depends on a large number of factors, some of which are

Labor costs and availability

Tax rate

Transportation (to bring in raw materials and ship out the products)

Public transportation (a poor system results in late arrivals and early departures by workers)

Rents

Pollution restrictions

Energy and water costs

For many companies, vacillating over whether to move or not, a small increase in the cost of energy can be the decisive factor. This extra cost may be the result of using less accessible natural resources, or it may be due to a requirement to switch equipment from natural gas to use low-sulfur oil to cut air pollution. After the company moves out, there is a snowballing effect: jobs decrease; welfare costs rise; taxes increase; and other industry decides to leave.

3. U.S. Energy Use

With only about 6% of the world's population, the United States accounts for half of the total use of natural gas and 30% of the oil use. The gross national product per capita is directly related to the energy use per capita; in both categories the United States surpasses any other nation by a factor of 2.[1] Energy has been among the fastest growing measures of U.S. life during the past two decades (Fig. 6-1).

---

[1] The relation between energy use and the standard of living is discussed for China over the past 1,000 years in the article by Earl Cook, Energy for Millenium Three, Technology Review, December 1972, pp. 16-23. The article is an interesting projection of the world's energy utilization well beyond the year 2000.

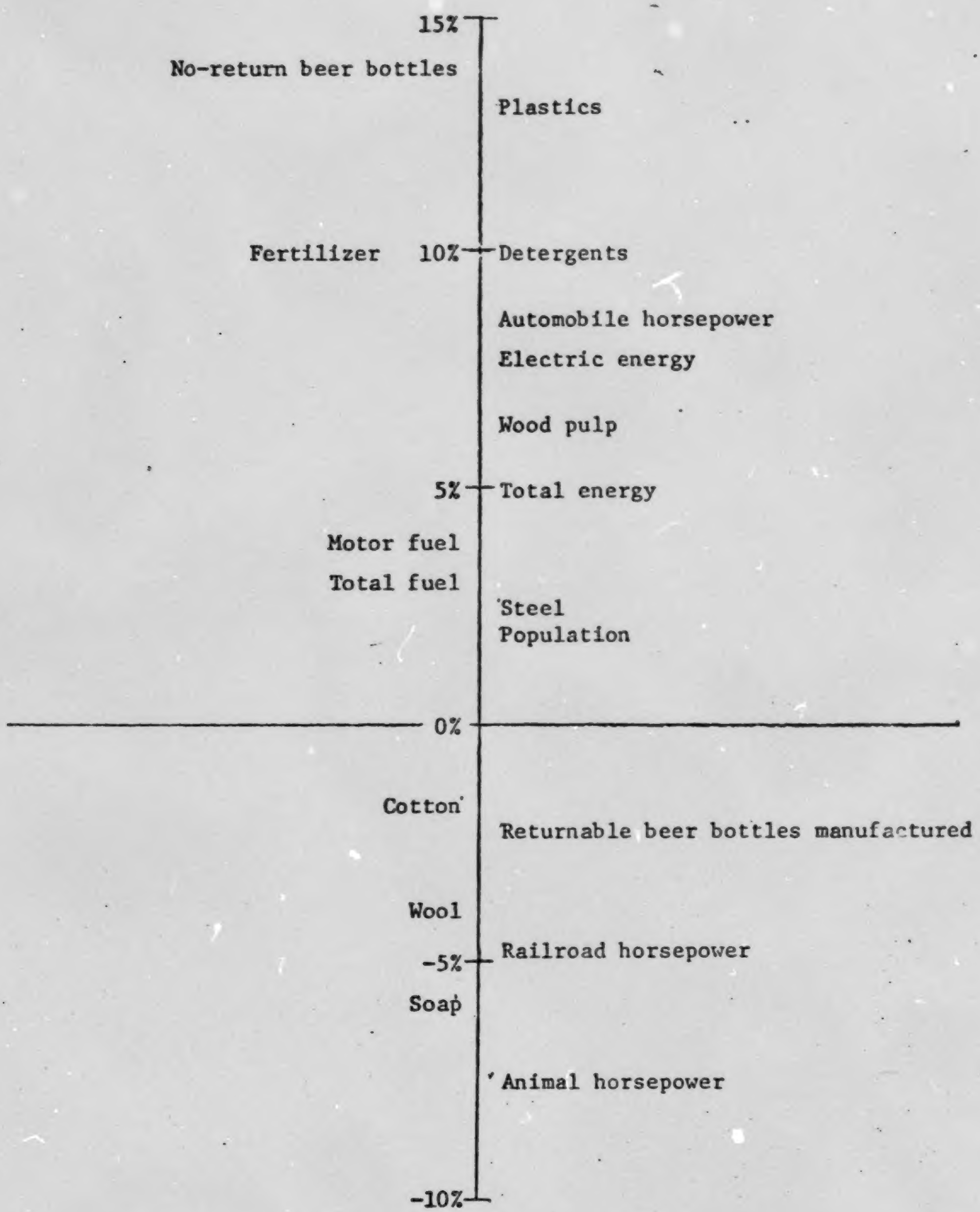


Fig. 6-1 Growth rates in the United States over the past two decades

United States energy consumption is equivalent to 400 people working full-time for each member of the population. In other terms, we average 40 horsepower per person[1] at all times. Of course, the use of energy is horrendously inefficient -- for example, we use a car capable of 280 horsepower to move one person. This is even more extravagant than the pharaohs of ancient Egypt, since it corresponds to 2800 slaves carrying one individual.

The energy is used in three general ways:

(1) Transportation -- 24%

(2) Industrial manufacturing -- 43%

(3) Home climate control and appliances, schools, government, commercial establishments -- 33%

The energy comes from the following sources:

Coal -- 20%	Petroleum -- 43%
Natural gas -- 33%	Hydroelectric -- 4%

Nuclear, geothermal, tidal, solar, wind -- less than 1%

From this general background, we turn next to a consideration of the detailed factors which have combined to lead to the current energy crisis.

---

[1] One horsepower was defined by James Watt about 1800 as the power which a strong horse can deliver. He measured this by having the horse lift a bucket of water from a well. One horsepower is equivalent to lifting 550 pounds one foot in one second, or about 3/4 kilowatt (0.746 kilowatt) in terms of electrical power. He must have had unusual horses, since an average work horse is able to furnish somewhat less than one horsepower.

The human being can provide one horsepower or slightly more for only a fraction of a second. Over reasonable periods of time, a man can furnish about 1/10 horsepower.

4. Contributory Factors

While the energy crisis is caused by the threatened change in the pattern of energy availability and use (Section 2), we can list a number of factors which have combined to bring us to the present situation.

(a) Environmental deterioration: The energy system impacts the environment at every step from the mining or drilling operations to the distribution and use. Coal mining is increasingly restricted to restore the natural beauty, particularly after strip mining. The damage from leakage in off-shore oil drilling has led to strict restrictions. The electric utilities are increasingly required to control air and thermal pollution. Overland electric power transmission is being shifted underground, and tankers for distributing oil must take stringent measures to minimize discharge into the sea. Steps such as these, while strongly in the public interest, exert pressure toward increasing energy costs.

(b) Accelerated growth: In the last few years, the use of energy has grown at an alarming rate. From 1880-1920, during the peak effects of the Industrial Revolution, energy use grew at a rate 60% higher than the growth rate of the gross national product (GNP). The major change during this period was the conversion to mechanization, which was energy intensive. This country then entered a period (1920-1965) during which major improvements were made in the efficiency of industrial processes, especially electricity generation. Over this period, energy growth was 20% slower than growth of the GNP.

Since 1965, energy use has been growing at a rate of 5% per year, well above the expected long-term growth rate of 3.4% per year for the GNP. This alarming growth rate is partly the result of such innovations as widespread air conditioning, but it is also partly due to a decreased efficiency in electricity generation. Many of the electric utilities have not anticipated the growth of demand or were blocked in building new facilities; the result is the prolonged use of outdated equipment with poor efficiency (and unnecessary contributions to air pollution). The situations in both Chicago and New York are critical; fortunately, the past few summers in both cities have been unusually free of prolonged hot spells.

(c) Electric energy growth: Closely related to the above, the growth of the use of electric energy is particularly impressive, now running about 8% per year (or doubling every 9 years). Today one quarter of the total energy used is consumed in the generation of electricity, and the fraction promises to hit one half by 1990.

In California alone, planning by the utility companies indicates a need for 130 new large generating plants by the year 2000. A Rand Corporation study has shown that, if the annual growth rate in use can be cut to 3%, the required number of new plants can be reduced to 23 -- still a difficult prospect since so little of the coastline is left which satisfies ecological, economic, and technological constraints. The question is: can the reduction from 8% to 3% possibly be achieved?

(d) Limited oil and gas reserves: As indicated earlier, the economically available reserves of natural gas and oil are limited in this country. The United States possesses 11 years supply of proven natural gas reserves, although large additional quantities could be obtained at higher cost from deep drilling or by coal gasification (in each case, the cost is estimated as four times the present level). In order to stretch the domestic supply, large quantities are now being imported, for instance from Algeria, and there are tentative plans to purchase from the USSR.

In oil, the United States now produces about 2/3 of its use, but production is only expected to increase slightly as use doubles over the next 15 years. Even if the controversial trans-Alaska pipeline is built, U.S. production will be increased by less than 20%.

In the world as a whole, convenient gas and oil supplies seem adequate to the year 2035, while sufficient coal is available for perhaps 300 more years.

(e) Impact on foreign relations: Thus, the United States depends heavily on imports of both oil and natural gas. The need for external sources of supply has to be closely inter-related with our foreign policy. In particular, the oil of both the Near East and South America (especially Venezuela) is important. Assurance of these imports becomes increasingly difficult as the oil-rich countries strive to improve their standards of living and emphasize their independence of the United States and Western Europe.

(f) Disappointments of nuclear energy: In the 1950's there were high hopes that by 1973 essentially all new electric generating capacity would be nuclear. Instead, we now find that only about 3% of the electrical energy comes from the 29 nuclear plants in operation. No full operating license has been granted in over a year, although 111 plants are either under construction or on order. With a 7-9 year lead time required for construction, it is now clear that by 1980 nuclear plants will still account for less than a quarter of our electric energy, even under the most optimistic assumptions.

The disappointment has arisen for several reasons. The hoped-for economy was often not realized as construction costs exceeded estimates. Nuclear plants are most economical when they can be operated continuously at normal power -- a primary reason for the common proposal that a nuclear plant be combined with a pumped-storage facility. The latter refers to a hydroelectric plant fed by an artificial lake at high altitude. During the night when customer usage drops, the nuclear plant output would be used to pump water up to the elevated reservoir. The next day, water from the elevated lake would be allowed to fall through turbines to generate electrical energy during periods of peak load. The total system would then operate the nuclear plant at maximum economy.

A second problem has been safety, particularly from an unusual accident or sabotage. During 1972 the Atomic Energy Commission recognized the possible safety problems and ordered reduced-power operation in several plants. A third problem has been thermal pollution when a river or shore region is used for cooling water. The questions on safety and thermal pollution have been the principal arguments of citizen groups organized to oppose new nuclear plants in several parts of the country.

In order to avoid lengthy public hearings centered on the safety and thermal pollution issues, the Public Service Electric and Gas Corp. of New Jersey is planning a nuclear plant 2.99 miles off the Jersey shore. The reactors will be on floating platforms roughly 400 feet on each side. The plant will be protected from high seas, hurricanes, tidal waves, and accidential ship collisions by two giant breakwaters, the one toward the ocean semi-circular and the one toward land straight, with a narrow channel between the two. In other words, a calm "lake" will be created surrounding the reactors.

The proposal seems to have such attractive advantages that there have been predictions of a chain of such plants around the U.S. coast. There is, of course, plenty of cooling water, and oceanographers feel confident that the thermal "pollution" will have no effect on marine life. In case of an accident, the reactors are well separated from populous areas by the three miles of ocean.

5. Alternatives or Options

If we accept the current situation as depicted in the preceding sections, the energy crisis is a fact. What options are available to ameleorate the situation? How can we avoid an increasing frequency of black-outs and energy shortages during the years ahead?

The possible actions divide into two general categories:

(1) Legal, political, and social

(2) Technological

In this socio-technological problem area, the technological solutions tend to be relatively long-term just because of the lead time required to build new facilities, even if the technical knowledge is available. In addition, the most promising technological options (as indicated in Section 6) are still in the research phase.

For the rest of this decade and probably well into the 1980's the United States will have to live with a continuous energy crisis, one which is likely to become steadily worse.[1] If we agree that available energy

---

[1] If the situation becomes sufficiently critical and detrimental to the quality of life, a national energy program similar to Apollo may be politically attractive. With an all-out effort, technology could yield remarkable improvements within a decade.

(particularly electrical energy) will not increase to meet the potential demand, we have to ask what steps can be taken to cut usage, or at least to slow down the rate of growth.[1]

Regulatory actions

The cost of electricity is controlled in each state by a Public Service Commission. The structure of rates is available to the government as a means to modify usage, and the Commission can place restrictions on utility advertising and operations.

At the present time, the rate structure is usually designed to encourage increased use: the unit cost decreases as use increases. Special rates are given for appliances, such as water heaters, which represent energy use during times when the total load is low (e.g., during the early morning hours).

To discourage use, the first step might be at least a flat rate structure. In addition, there could be a penalty for use during the hours of peak load, although this change would not be simple to do technically.

There are objections to such a change in policy. In New York City, as an example, the two largest users of electricity are the subway system and the local government (for lights, building operation, and so forth). Penalizing these customers would directly lead to higher subway fares and reduced mobility for the poor, higher taxes and a further exodus of the middle class. Other large users are industry and apartment houses; in neither case are we anxious to increase economic pressures.

---

1. Two articles on energy conservation and reduction of needs appeared in Science December 8, 1972 (pp. 1079-1081) and December 15, 1972 (pp. 1186-1188).

A second difficulty of using the regulatory mechanism to decrease usage is that the elasticity of demand is very low. A price increase has little effect on sales. Most people just do not associate costs with specific appliances. The typical homeowner has a poor picture of the annual cost of each of his uses of electricity (television, range, washing machine, lights, and so on). The normal family is quite willing to limit its use of air conditioning when the outside temperature is 80°F, but much less enthusiastic when it is 98°F. Unfortunately, the latter condition is when the energy crisis is critical.

While regulatory authority can readily be used to attempt to reduce energy usage, there is little likelihood that the effects will be very large. When the energy crisis is as severe as it promises to be in the next few years, every helpful measure should be taken, even if the benefit is small.

Education and advertising

As a nation, we are wasteful of energy -- not only in the automobiles and similar extravagant uses of power, but also in many smaller ways. Few buyers of air conditioners worry about the efficiency of the unit or whether it is unnecessarily powerful for the planned use. During the hot summer days, rooms are often kept appreciably cooler than would be normal during the winter.

Home insulation is often inadequate and incomplete, with the result that large amounts of heat are lost during the winter. While oil or gas heat should permit efficiencies as high as 70%, the typical home furnace operates at 40% simply because of maladjustment. Lights are left on unnecessarily; in some, modern office buildings a single switch controls an entire floor, so that one worker uses illumination sufficient for hundreds. In certain cases, lights can not be turned off without disconnecting the entire electrical system. Incandescent lamps are still much more common than the more efficient fluorescent lighting. The frost-free feature of refrigerators increases energy consumption.

In proposals to cut the growth rate of electric-energy usage from its present 8% per year to something like 3%, the principal hope is to reduce the energy waste in homes and commercial establishments. Under pressure from the regulatory agencies, some of the electric utilities have now converted their advertising toward a public education program to save energy. The company, which not long ago advertised to stop burglary by leaving a light burning, now publishes long lists of ways to save energy -- such detailed instructions as

Don't put hot foods in the refrigerator
Vacuum only on weekends
Keep your oven closed while roasting; each peek costs 25°F
Use the dishwasher once a day
Leave the thermostat at one setting
Do all cooking for a day's meals at one time
Keep blinds and windows shut on hot summer days
Reduce the wattage of corridor and closet bulbs

There is a real question whether such a program can make a significant difference in the use of electric energy. The problem is again the "tragedy of the commons" discussed on page 178. Each person has to be willing to sacrifice some personal benefit for the overall good; he can not act to maximize his own pleasure or comfort.

The problem is also complicated by the uncertainties of how to present an effective educational program.

Legislation

Energy use can also be reduced by legislative mandates. Slight improvements can sometimes be made relatively painlessly, for example by staggering working hours to reduce the peak loads on mass transit. Wastage can be taxed, although not always easily. If an electric utility is taxed for using low-efficiency equipment, the cost of electricity to the public rises even more (it probably is already high because low-efficiency generation tends to use outdated machinery which is expensive to operate). Tax on waste might be possible by making the tax on an appliance such as an air conditioner depend on the efficiency of the unit. Certainly the manufacturer can be required to display the efficiency of the device, hopefully in terms understandable to the public.[1]

A significant reduction in energy use could be accomplished over a very long period of time if building codes were changed to require insulation in homes and heat-reflecting glass on commercial buildings. Estimates are that 40% of the energy used in interior climate control could be saved in all new construction.

More drastic legal measures have been suggested. Electric home heating might be forbidden (electric generation is only about 30% efficient, and often overlooked is the cost of the air pollution which results from the generation by today's plants which use primarily fossil fuels). For the same reasons, the electric automobile seems to promise no real hope in the next two decades.

---

[1] Room air conditioners vary almost 3 to 1 in efficiency, from 4.7 to 12.2 Btu of cooling per watt-hour of electrical energy. New York City now requires that this information be displayed.

If the energy crisis becomes sufficiently severe, even more drastic legislation may be proposed. The manufacture or sale of air conditioners and appliances could be forbidden or limited, perhaps by an excessive tax or by using a lottery to determine each year's lucky purchasers. When such thoughts had been presented in testimony before a Congressional committee studying the energy crisis, a Congressman responded, "Have you no ideas that are useful to this committee?"

Final comment

In this section, we have looked briefly at some of the measures which might be taken to ease the energy crisis during the next decade. Unless we can imagine a radical change in the American pattern of living, no combination of these steps is likely to provide more than a short-term amelioration. Over the long term, new technology must provide the solution, and hopefully before the impact of the energy crisis is disastrously disruptive.

6. New Technology

Table 6-1 shows the rapid growth of the federal program in research and development in energy problems. During a four-year period when science and engineering research programs were totally at nearly a constant dollar level, the energy program has nearly doubled from $363 million to $622 million.

The table is important because it indicates the directions judged to be most important by the federal officials and the advisors who have guided this growth. For example, the fifth row from the bottom (Geothermal

TABLE 6-1

Federal Funding of Energy Research and Development
(in millions of dollars)

	FY 70	FY 71	FY 72	FY 73
Coal Resources Development				
Production and Utilization R&D,	$13.2 M	$15.4 M	$14.7 M	$19.[illegible] M
incl. gasification, liquifaction and MHD	13.5	18.8	31.1	45.3
Mining Health and Safety Research	3.7	14.8	31.0	30.1
Petroleum and Natural Gas				
Petroleum Extraction Technology	2.7	2.7	3.2	3.1
Nuclear Gas Stimulation[b]	3.7	6.1	7.0	7.5
Oil Shale	2.4	2.7	2.6	2.5
Continental Shelf Mapping	-	-	5.0	7.0
	6.0	6.0	6.0	6.0
Nuclear Fission				
LMFBR[b]	144.3	167.9	236.6	259.9
	-	-	0.8	1.6
Other Civilian Nuclear Power[b]	109.1	97.7	90.7	94.8
Nuclear Fusion				
Magnetic Confinement[b]	34.3	32.3	33.2	40.3
Laser-Pellet[b, c]	3.2	9.3	14.0	25.1
		-		
Energy Conversion with Less Environmental Impact				
Cleaner Fuels R&D-Stationary Sources	19.8	17.4	24.5	29.5
SO_x Removal	-	-	2.6	15.2
Improved Energy Systems	0.8	3.0	2.4	2.8
Thermal Effects R&D	0.8	0.6	0.7	1.0
	1.5	1.8	3.2	6.8[d]
General Energy R&D				
Energy Resources Research[c]	1.1	5.0	9.8	13.4
Geothermal Resources	0.2	0.2	0.7	2.5
Engineering Energetics Research	2.9	2.7	4.0	4.7
Underground Transmission	-	0.8	0.9	1.0
Cryogenic Generation	-	-	-	1.0
Non-Nuclear Energy R&D	-	-	-	1.5
	$363.2 M	$405.2 M	$524.7M	$621.6 M

Resources) shows a very rapid growth from $700 thousand to $2.5 million this year, even though the total effort is still small. The growth demonstrates the promise of this potential source of useful energy, the low level shows that there has been very little attention to this possibility in the past.

In this section, we discuss briefly a few of the important and exciting directions of energy research and development.

Geothermal energy

Pacific Gas and Electric Company now is operating an electric generating station from the geysers in the Sonoma Valley of California. The energy is drawn from the heat well below the surface of the earth. This concept of using this fantastically large supply of energy is not new: for almost 70 years, electricity has been generated from the steam coming out of the ground at Larderello, Italy, and buildings in Reykjavik, Iceland are heated by hot water from the geyser fields.

Hot springs and geysers are exceedingly rare. At any place, however, if one drills far enough below the surface of the earth, he hits hot, dry rock -- often at 900°F. In many places, this source of energy is near enough the surface (as little as one mile) to be reached by economical drilling. These regions often can be found by the large number of very small local earthquakes.

At Los Alamos, New Mexico, for example, two 15,000-foot holes are being drilled down to the granite which is at 600°F. The rock will be cracked so water can penetrate. Water is then allowed to fall into one hole, it turns to steam at the rock, and steam rises to the surface through the second hole. This steam is used for the generation of electricity.

Superficially, the scheme looks exceedingly attractive: steam is obtained for only the cost of two holes. The energy is free.

Nature is seldom so kind, and it is not surprising that there are technological difficulties. The steam which rises is at a temperature much lower than needed for efficient generation of electricity. The rising steam may have excessive amounts of sulfur and salts, with air pollution the result unless costly equipment is added. Because of uncertainty about the economics of clean electrical generation, the future place of geothermal sources in the total energy picture is not clear.

Extraction of fossil fuels

The first three lines of Table 6-1 refer to the federal programs directed toward better utilization of the enormous supplies of coal available within this country. As shown there, as little as four years ago, very few research-development dollars were allocated to this field; it is only very recently that planners have recognized that the limited reserves of gas and oil can be conserved by a switch to coal in many applications. The result is a rapidly expanding program in the Department of the Interior under two organizations: the Bureau of Mines and the Office of Coal Research.

A major thrust of this program is toward clean coal gasification: the production of clean gas to replace natural gas. Principal limitations are the cost of the gas produced and the large quantities of water needed for the process at a time when fresh water is also in seriously short supply in much of the country.

One interesting item in Table 6-1 is the third line for mining health and safety research. 200 miners are killed each year and many more develop

Black Lung disease. In the latter category, the government now spends one billion dollars annually for sufferers and survivors, seemingly disproportionate to the 30 million devoted to prevention and to improvement of conditions in the mines.

Solar energy

For many years there have been hopes that energy from the sun could be used directly for heating and particularly for the generation of electricity. In Table 6-1, solar-energy research is included within the line entitled Energy Resources Research, sponsored by the National Science Foundation.

Use of solar energy for electrical generation is feasible but economically unattractive. In the past, cost determinations have shown a figure as much as 100 times that for fossil-fuel systems. To collect a significant amount of energy, the plant would have to be located in a cloudless desert area. The collector would be a carefully designed surface, perhaps almost two miles on each side, and made of very thin layers of materials chosen so that it absorbs most of the heat but radiates very little away.[1]

Even if enough energy can be collected, there is a problem because electrical generation is most efficient when the turbines are driven by steam at high temperatures. Even the most optimistic plans for solar generators envision a relatively low efficiency process, which is economically attractive only if the lost heat can be used for other purposes such as water desalinization. Thus, the feasibility of the direct use of solar energy for electrical generation is still very questionable, although there are a few engineers who are ardent advocates.

---

[1] Unfortunately, black bodies which are good absorbers also radiate well. The solar energy is primarily in the visible part of the spectrum, while most of the radiation from the collector is infrared. The trick is to build a surface which looks black to visible light, but like a mirror to infrared. This can be accomplished by layers of thin films: the outer layer absorbs visible light but is transparent to infrared; the inner layer reflects the infrared. One further problem is the high costs of keeping such a surface in a vacuum so air doesn't carry away the heat, and dirt subtract from the collection efficiency.

Solar energy for home heating and air conditioning, on the other hand, does seem attractive; it is already economical in such areas as southern California (outside the smog region), and cheaper than electric heat in the New York vicinity, although still twice the cost of oil or gas. Under research since 1938 at M.I.T., the system uses a roof collector; water passing by is heated and then stored. The most promising arrangement uses electric heat to complement the solar energy, since the latter demands sunshine on at least every other day.

Other ideas

There are several other attractive research areas. A total energy system for the home is being used in newer apartment complexes. Here an internal combustion engine drives a generator for on-site electrical energy. The exhaust gases and waste energy are used to heat or air condition. The same system can also handle solid waste compaction (and, indeed, in a few cities waste burning in incinerators is now being used to give useful energy).

In a parallel experiment with self-sufficient systems, fuel cells are being used in a few homes. With hydrogen a possibility as an energy source, fuel cells yield high efficiency (50-55%, compared to fossil-fuel generation at 30-35%) and very little pollution (water and carbon dioxide are products, and thermal pollution is reduced by the high efficiency). In both this application and the above total home energy system, the initial equipment cost and maintenance problems are severe; probably the homeowner would rent the equipment from a utility company.

Table 6-1 indicates the major research in the liquid-metal, fast breeder reactor (LMFBR), in the long-term hope of controlling nuclear fusion, and in generation and transmission using cryogenics (temperatures very near absolute zero, so that there is no resistance to electricity and hence no energy lost in heat).

In this field of energy, there is also an assortment of much more radical proposals, often promoted enthusiastically by individual scientists and engineers. The tides contain enormous amounts of energy which is unused. Winds provided energy years ago when windmills dotted the western European landscape. In the ocean, there are areas where the temperature changes rapidly and heat energy might be extracted. Solar energy might be used to grow plants to be burned or to produce chemical changes which could store energy. Finally, there is a proposal to orbit an enormous solar-energy collector. The energy would be converted to microwave frequencies and beamed on to an earth receiving station.

7. Final Comment

The energy crisis is so very important because the elements of extreme social difficulty already exist. Natural gas, in particular, is limited. Electrical generating capacity is only slightly greater than demand in peak periods, while electrical use is growing at the alarming rate of 8% per year.

The construction of new generating capacity or the conversion between fuels involves long time lags, often seven years. Promising new technology will not be significant for an even longer period of time. Thus, the energy crisis is almost sure to worsen from now well into the 1980's. Severe legal restraints on energy use will be necessary to minimize the disruption of the social system.

CHAPTER 7

NOISE ENVIRONMENT

The spoken word "see" is recorded and played over and over again to a listener. If the word is repeated twice each second for three minutes, a total of 360 times, the average young adult hears something like six different words, with the changes occurring possibly 30 times. He will hear such related forms as "ease", "fee", "tee", and so forth.[1]

The astonishing feature of this experiment, however, is that the results are age-dependent. At the age of 5, the child experiences no illusions; by 8, he hears the full group; this "capability" lasts until the 20's, when it starts to decline; by age 65, he is back to the status of the five-year-old child. Apparently the perception of illusions depends on the way in which the individual's nerve system and brain process the sound inputs.[2]

The experiment illustrates the fundamental problem when we try to study the effects of various sounds on human beings. A sound signal is a real, measurable thing: it is simply a change of the pressure in the air. When it reaches a human being, the sound signal starts some of the 30,000 hairs in the inner ear vibrating -- which hairs vibrate and how much is determined by the particular sound. As these hairs move, nerve cells at the base of the hairs are actuated, and electrical signals then travel from these cells to the brain.

---

[1] Richard M. and Roslyn P. Warren, Auditory Illusions and Confusions, Scientific American, Dec. 1970, pp. 30-36.

[2] One immediately thinks of using such tests to measure verbal capabilities of children in the lower elementary grades or, at the other end of the spectrum, to obtain an indicator of aging.

Thus far, everything is straightforward and reasonably understandable (even though we do not understand as yet the coding or system that determines which nerves are energized by particular sounds). The human reaction to these electrical nerve signals reaching the brain is, however, vastly more complicated. A particular sound may cause many different reactions from different people. Responses from the same person may differ greatly depending on what he is doing at the time, his fatigue, his state of mind, and so forth.

Thus, the study of the effects of sound on people is an exceedingly complex subject, which brings together the physics of sound (called acoustics) with psychology to form the subject of psychoacoustics.

In this chapter, we want to look at one type of sound: noise, which is undesired and usually unpleasant sound. Specifically, we consider how noise can be described quantitatively, the effects on people, and then the sources of noise and how they can be controlled. In other terms, the goal is the study of the noise environment, with emphasis on the alternatives available to us to control this form of pollution.

1. Measuring Noise

Noise, or more generally sound, is a variation of the air pressure. At the surface of the earth at sea level, the normal air pressure is approximately 14.7 pounds per square inch. The air surrounding the earth exerts a pressure of 14.7 pounds on each square inch of the surface, or more than 2100 pounds on each square foot.

When a speech signal is sent out from a speaker's mouth toward a listener, the talker forces the air pressure at his mouth to vary in a way corresponding to the sounds being sent. This signal or pressure change travels out in all directions at a rate of approximately 1100 feet/second or 600 miles/hour -- the speed of sound in air.[1] The pressure changes reach the listener's ear, where they cause the hair vibrations mentioned earlier.

This pressure wave, travelling out from the speaker until it reaches the listener, is characterized by two features: the size of the variation which is called the amplitude, and the pitch or frequency -- the number of cycles of variation each second. Thus, amplitude and frequency are two basic measures of the characteristics of noise.

2. Frequency

When a tuning fork for middle A is struck, the sound emitted is a pure musical note, in this case 440 cycles/second. The pressure resulting from this note changes with time as shown in Fig. 7-1. The pressure increases and decreases, as time progresses, around an average value (14.7 pounds/square inch). Each cycle lasts 1/440 second, or there are 440 cycles per second.

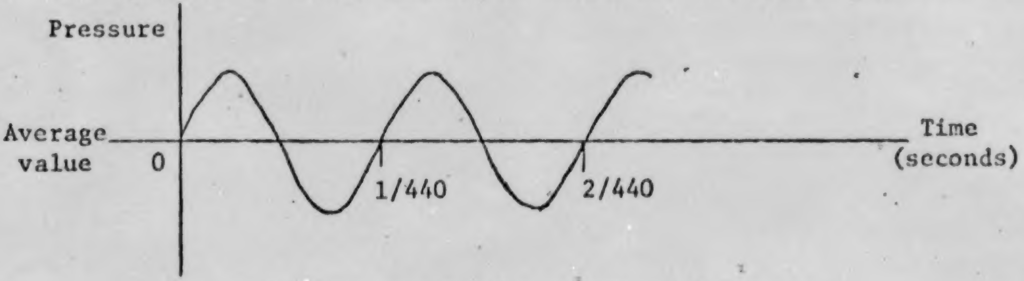


Fig. 7-1 Pressure in the presence of a pure middle A note.

---

[1] This speed of sound in air is the basis for calculating the distance to a lightening stroke in a thunderstorm. Light travels almost instantaneously (186,000 miles/second , the sound of thunder at the speed of sound -- 600 miles/hour or 10 miles/minute or 1 mile every 6 seconds. Hence, each 6 seconds between the flash and the clap correspond to a distance of one mile.

In mathematical terms, the pressure is varying sinusoidally at a frequency of 440 cycles/second.

When middle A is played on a flute, the sound that results contains not only a variation at 440 cycles/sec, but also the second, third, fourth and fifth harmonics at 880, 1320, 1760, and 2200 cycles/second, respectively. Each musical instrument has a unique group of harmonics, so that we can recognize the instrument by listening to a single note.

A fundamental theorem (the Fourier theorem) states that any sound signal can be represented as the sum of pure notes or sinusoidal signals.[1] Any noise is simply the adding together of pure notes or frequency components. Most noise signals, to be sure, consist of thousands of separate frequencies, but the important thing is that this decomposition does exist. As a result, we can describe a noise signal in terms of its frequency components.

This frequency decomposition of sound is particularly important in our study of noise pollution because the human ear responds to only a relatively narrow range of frequencies -- from about 30 cycles/sec[2] to 18,000 cycles/sec in the teenager with ideal hearing, and an appreciably narrower band for older people. The only noise which affects people is that part within this frequency band. (The whistles used to call dogs usually have a frequency around 20,000 cycles/second, which can not be heard by human beings).

---

[1] See ECCP, "Man and His Technology," McGraw-Hill Book Co., New York, N.Y., 1973, Chap. 6.

[2] To measure frequency, we are using cycles per second throughout this chapter. About a decade ago the engineers and scientists decided that a new name should be given to this unit, and it is now officially called Hertz, in honor of Heinrich Rudolf Hertz, the German physicist (1857-1894) who first demonstrated the production and reception of radio waves.

The sensitivity of human hearing varies markedly with the frequency of the sounds. Figure 7-2 shows the weakest signal the average, normal person can hear as the frequency changes from 32 to 16,000 cycles/second (almost four octaves below middle A to five octaves above, with each octave increase meaning a multiplication of the frequency by 2). The ear is most sensitive at approximately 4,000 cycles/second. The curve varies rather smoothly toward a minimum between 2,000 and 4,000; the bulge downward near 4000 arises because in this frequency range the sound wave striking the face tends to travel around the head to the ear (through the process called diffraction in physics).

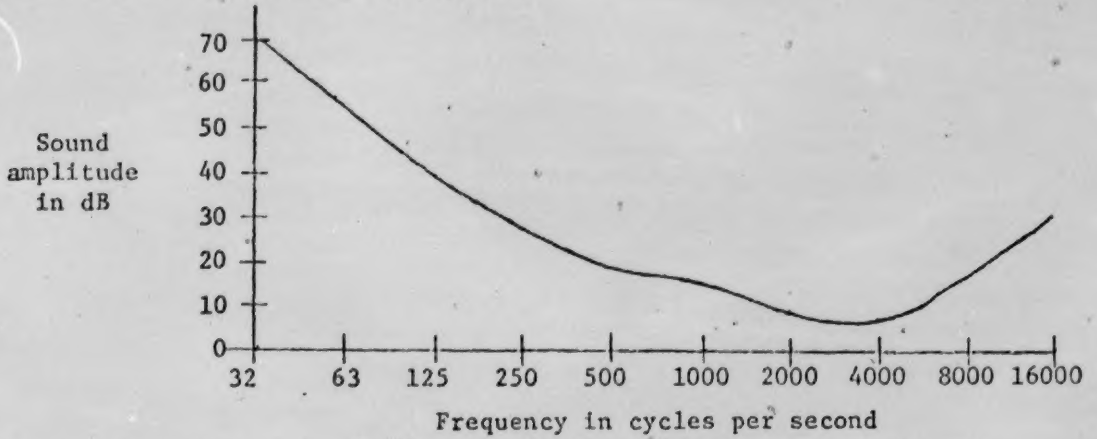


Fig. 7-2 Weakest signal the average person can hear at various frequencies (the meaning of the amplitude scale is explained in the next section)

The curve of Fig. 7-2 is measured with an audiometer -- a device to measure a person's hearing capability. In the earliest hearing measurements during the early 1800's, a tester simply spoke at a constant amplitude as he moved away from the subject. Even in the 1930's the author's hearing was checked in elementary school by essentially the same procedure.

By the mid-1800's a tuning fork was used to generate a pure note, with the distance between the fork and ear varied. Electronic equipment became available about 1920, after the invention of the vacuum tube permitting precisely controlled amplification of the electronic signal. The great advance in audiometers came, however, at the close of World War II, motivated by the desire of the Army and Veterans' Administration to measure accurately the hearing loss sustained during combat. In recent years, speech at a carefully controlled amplitude has been used in addition to the pure-tone tests in audiometry equipment because certain hearing problems do not appear when the subject listens to only the pure tones.

Figure 7-2 shows that noise pollution which is most annoying to people will be that in which there are large components in the frequency range from 2,000 to 4,000 cycles/second. Here the psychological aspects have to be introduced as well. When subjects are given different types of noise and asked to adjust their levels or amplitudes so that all sound "equally as annoying," we find that there is a significantly more intense dislike of noise from 2,000 to 8,000 cycles/second than would be predicted from Fig. 7-2. Not only is human hearing more sensitive at certain frequencies, but in addition a person finds this range of frequencies particularly unpleasant. Noise in the band 2,000 - 8,000 cycles/sec must be especially minimized if we are to improve the noise environment.

3. Amplitude

The second important characteristic of a sound signal is the amplitude or, in general terms, the loudness. At 4,000 cycles/second, the weakest sound anyone can hear represents a pressure variation of 2 x 10^{-9} times atmospheric pressure (i.e., two billionths of atmospheric pressure). The loudest sound the ear can tolerate without pain is about 10 million times as great, or 2/100 of atmospheric pressure.[1] The human ear can detect a fantastic range of sound amplitudes.

In Section 7 of Chapter 2, we plotted curves of exponential growth covering an enormous range. There we found it convenient to let each vertical division represent the same multiplying factor. That is, each time we moved up on the graph by one division, we multiplied the quantity by 2 or any convenient factor. It is useful to measure sound amplitude in the same general way. To do this, we arbitrarily call the weakest detectable signal 0 dB (the measure we are using has the name dB, which stands for deciBels).[2] Every time we multiply the sound amplitude by 10, we add 20 dB. Thus, compared to the weakest detectable signal,

a signal 10 times larger is	20 dB
a signal 100 times larger is	40 dB
a signal 1,000 times larger is	60 dB
a signal 10,000,000 times larger is	140 dB (the signal at the pain level)

This use of a dB amplitude scale is, first of all, a convenience. We don't have to talk about signals 2,000,000 times as big as others; in other words, we avoid the use of the multi-digit numbers. Also, the human ear responds in approximately this way. The increase in loudness as noise goes from 20 to 40 dB is about the same as when it goes from 80 to 100 dB.

---

[1] At an Apollo or large space rocket blast-off, spectators are kept far enough away to ensure that the sound does not exceed this maximum level. 2/100 of atmospheric pressure is about 0.3 pounds/square inch, or a pressure of 60 pounds on a man's chest and stomach. Thus, the statement that the awed spectator makes that he can feel the pressure waves on his body is not unreasonable.

[2] The original unit was in Bels, named after Alexander Graham Bell, with the second L somehow being lost. A more convenient unit is tenths of a Bel or a deciBel.

A sound or noise signal is measured in deciBels or dB. Sounds can appear in the range from 0 dB to 140 dB. The vertical scale of Fig. 7-2 is now understandable. At 50 cycles/second, the signal has to be 60 dB to be heard by the average person -- it would have to be 60 dB or 1,000 times larger than if the frequency were 4,000 cycles/second.

Unfortunately, when we study noise, we are really interested not in the amplitude of the sound, but rather in how loud it sounds to a human being. As Fig. 7-2 shows, sounds around 4,000 cycles/sec seem much louder than those at either low or high frequencies. Usually noise signals are made up of many different frequencies. How do we account for the relative amplitudes of the different frequency components and also the sensitivity of the human ear?

The simplest answer is a new measure known as the dB(A) scale. When a noise meter measures in terms of dB(A), the different frequency components of the noise are weighted according to the sensitivity of the ear (the parts of the noise around 4,000 cycles/second given the greatest weight). The meter then sums the effects of the different frequencies to read an apparent loudness value -- say 85 dB(A). Thus, when we say the noise level is 85 dB(A), this means that the human ear perceives this noise to be about the same loudness as a pure, 4,000-cycles/second note at an amplitude of 85 dB. The dB(A) scale is an attempt to measure the apparent loudness to a human listener.

There are a large number of other loudness scales which are used in noise studies. The perceived noise level (PNL) was developed particularly to measure aircraft noise near an airport. There are also dB(B); db(C), and dB(D) scales, noise criteria (NC) curves, effective perceived noise level (EPNL), traffic noise index (TNI), and so forth. Many of these have been developed for particular measurements, such as factory noise, street noise, etc. They often involve studies with sample groups of subjects to determine when people feel subjectively that two noises are equally annoying.

Most of the inexpensive noise meters measure dB(A), and this is the scale most often quoted in the newspapers and public media. Regardless of the particular scale used, the results are generally comparable and, for the various noise situations shown in Fig. 7-3, we obtain approximately the measurements shown.

4. Other Characteristics of Noise

The effects of noise on people depend not only on the frequency and amplitude, but also on the duration. A very brief noise, such as a gun firing, has much less effect than a long-duration noise of smaller amplitude.

The importance of this characteristic was recognized as early as 1936 in the Walsh-Healey Public Contract Act. The U.S. Department of Labor has the responsibility of setting noise standards for the protection of workers at any manufacturing plant which has a federal contract of more than $10,000. It was 1969 before an attempt was made to enforce this worker-protection provision, and today monitoring tends to be irregular and occasional.

Under these regulations, the allowable noise levels are shown in Fig. 7-4. 90 dB(A) is the limit for exposure during an 8-hour day; 115 dB(A)

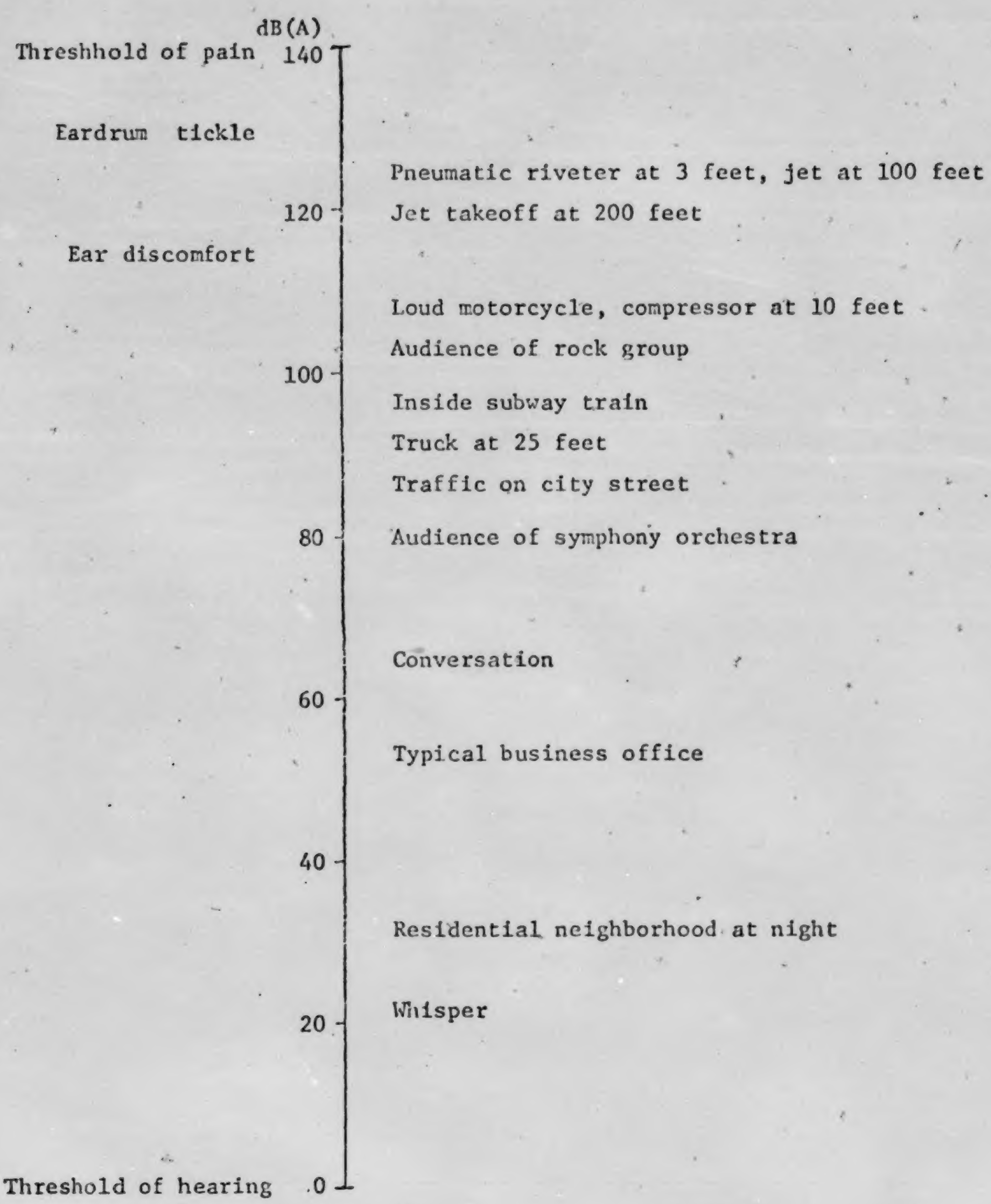


Fig. 7-3 Typical noise levels

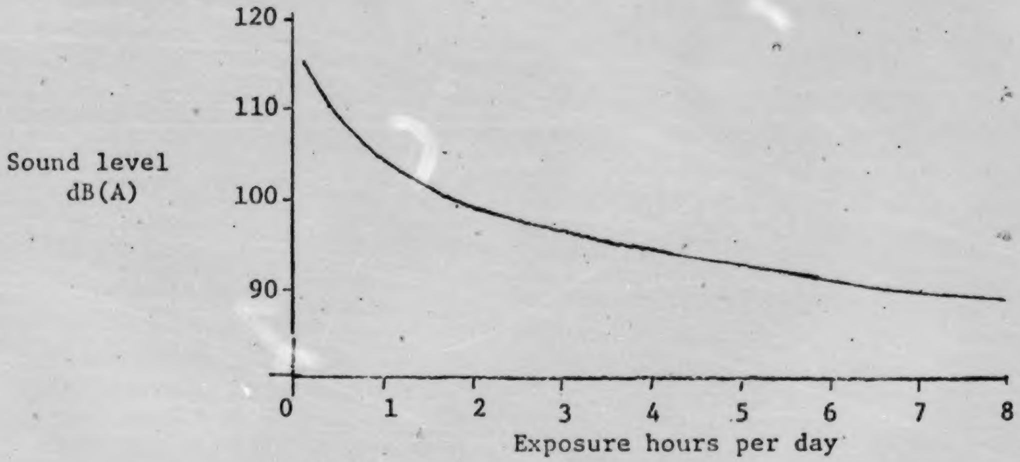


Fig. 7-4 Allowable noise exposure in industry. If there are two or more exposures at different levels, the limiting value is determined by the fraction of the allowable time at each level. This total of these fractions should not exceed unity. For example, 5 hours at 92 dB(A) (where 6 are allowed) contributes 0.83; 1 hour at 95 dB(A) (when 4 are allowed) contributes 0.25. The sum is 1.08, which exceeds the allowable value of 1.

when the duration is only 1/4 hour. In addition, impulsive or impact noise (lasting for a very brief interval) must not exceed 140 dB(A). In general these levels have been set from past histories which show that 1/6 of the people working in an environment at this level will have severe hearing problems in later life, compared to the population average of 1 in 14. The "allowable" levels are not what one would term "safe."

If these limits are exceeded, the manufacturer must make every reasonable effort to reduce the noise to which workers are exposed. If the noise level can not be reduced below the Walsh-Healey limits, workers must be given earmuffs or ear plugs for protection.

There are other characteristics of noise which we know from experience influence the effect on people. The unexpectedness of a noise is certainly important. The human being has the capability of some sort of self-blocking of loud noises. When a balloon is popped in a small room, if we anticipate the noise, it is not particularly disturbing. Apparently the person activates some sort of a mechanism which either partially blocks nerve signals to the brain or limits brain processing of these signals.

Human hearing also has remarkable capabilities to respond to certain sounds while rejecting others. The most famous example is called the "cocktail party effect." In a room with a very high sound level, a person can understand a single conversation in which he is particularly interested -- even though on a scientific basis that conversation should be totally swamped by the general noise.

Other largely unexplained phenomena are found in a study of human hearing. There are masking effects, in which the simultaneous appearance of two different sounds may result in one being totally unperceived by the listener. Certain sounds seem to be particularly annoying to some people, while almost unnoticed by others. A familiar example is the shrieking sound as chalk is pulled across a blackboard.

A classic example of the complexity of human hearing is the confirmed ability of a sleeping mother to awaken rapidly when her baby utters a weak call for help, while she sleeps soundly through enormously louder sounds from outside traffic, thunderstorms, and the like.

All these unexplained phenomena combine to make the study of human hearing, and of human senses more generally, a fascinating frontier for engineering research. At the same time, the unknowns complicate the study of noise pollution. Since we really don't understand how a person hears, the determination of the effects of noise pollution on people is exceedingly difficult.

5. Physiological Effects of Noise

The average noise level in parts of New York City during a working day is approximately 80 dB(A) and seems to be increasing at the rate of 1 dB(A) per year. Comparison with the allowable industrial levels of Fig. 7-4 shows that in one decade the city noise will exceed the value allowed for 8-hour exposure. These data have led some environmentalists to predict that by the end of the century half of the residents of New York City will have serious hearing problems.[1] The situation may be worse than this, since there is the possibility that long-term exposure to the Walsh-Healey allowable levels may cause hearing deterioration.

While hearing loss from noise pollution was first recognized in 1830, concern is reinforced by several studies in the past few years. In the United States, hearing seems to deteriorate rather generally with age. Not only is there some reduction of sensitivity, but there is normally a marked decrease in the maximum frequency which can be detected. The highest key on a piano is about 4,000 cycles/second; it is not unusual for older people to be unable to hear this sound. In direct contrast, studies of African natives, who have lived in a quiet environment, reveal no

---

1. This is clearly not a valid prediction. Once a large fraction of the population is deaf, these people will be severely limited in activity, and the city noise level will fall. The growth of deafness will not be exponential, but will follow the sigmoid characteristic of Section 6 of Chapter 2.

degradation of hearing with age. One might argue that the small life expectancy means only the hardiest survive, but there seems no logical reason why the "hardiest" should have exceptional hearing.

In another experiment, four years ago a quarter of the entering freshmen at the University of Tennessee had hearing deficiencies. A guinea pig was outfitted with one ear plugged, the other open. The animal was exposed for 88 hours over three months to rock music adjusted to the sound level popular with teenagers. This exposure corresponded to a typical experience of a high school student. At the end of the test period, the plugged ear of the guinea pig retained excellent hearing, but in the open ear most of the nerve cells had been destroyed. While it certainly is true that results for guinea pigs can not be carried over to human beings, experiments such as the above have raised serious questions about the effects of loud music and the general sound environment.

There is a scarcity of medical data on which to base broad studies of the changes in hearing capabilities of the population. Hundreds of thousands of Americans have been given batteries of medical tests including audiometry or hearing, particularly in the automated multiphasic health testing centers which have become so popular in the last few years. Unfortunately, these data are not easily recoverable, and they also are complicated by the use of different equipment in separate locations.

Measures of deafness or hearing deficiency are only rough estimates. It does appear that approximately 10 million Americans now have hearing problems, although perhaps only 10% of these are deaf. Somewhere between 6 and 16 million workers are currently exposed to industrial noise levels exceeding the Walsh-Healey standards.

The latest research hints that the situation may be even more serious. Studies at Mt. Sinai Medical School in New York have revealed that hearing ability falls off much more sharply with age than measured by the usual audiometry tests.[1] The older person is less able to understand either speech in background noise, or speech which is interrupted by having segments sliced out. It is not clear whether this loss of ability comes from the hearing mechanism in the ear. It may be just a decreased capability of information processing in the brain, since the human being has a remarkable ability to listen to large blocks of words. In the days when the telegraph was a vital communication medium, the highly trained operator receiving a message would write as much as 12 words behind the signals he was hearing. This lag allowed correction for errors. In the same way, an expert typist is aware of 4-5 words ahead of the one being typed. Very possibly this ability to hear or read large groups of words at a time diminishes with age.

Thus far in this section we have discussed a single physiological effect of the noise environment: poorer hearing. There are other observed effects which are even less understood. Noise causes:

Fatigue
Increased heart rate
Constriction of blood vessels
Dilation of brain blood vessels
Gastric activity
Muscle tension
Sweating
Dilation of pupil of eye
Lower electrical skin resistance

---

[1] In an audiometry test, the subject listens through an earphone to a pure note which is gradually changed in amplitude until it can just be detected. After the test is repeated with frequencies over the normal hearing range, we obtain a picture of hearing sensitivity similar to Fig. 7-2.

There is a suspicion that a sustained noise environment may be related to heart and neurological problems. Very loud noises have induced air bubbles in the veins of small animals, with resulting death. Recent research has suggested that excessive noise may affect the fetus before birth, possibly even contributing to malformation. Finally, the claim has been made that the anal sphincter muscle resonates at about 75 cycles/second; a sustained sound environment at that frequency and large amplitude causes muscular oscillation.

Apparently some people have the ability to adapt to a noise environment, and the effects decrease with time. Others show no adaptation, and a third group becomes more disturbed as time passes.

As we look at physiological effects, we find very large areas of uncertainty. This difficulty arises again and again in environmental problems. Scientific knowledge is just not adequate to predict the effects on human beings, whether we are discussing noise or air pollution, crowdedness, and so forth. The worry is that the physiological effects will be irreversible: before we obtain the scientific knowledge needed to make intelligent decisions, permanent changes will have occurred.

6. Psychological Effects

If the physiological effects of noise are uncertain, psychological effects represent a much worse situation of ignorance. Very few facts are available.

Experiments at Stanford University have shown that outside noises during the night can disturb deep sleep without actually awakening the subject. If the noises are repeated every time the man falls into a deep sleep (as measured by brain waves), he is unusually tired when he awakens in the morning. Furthermore, dreaming can be prevented, with the resulting possibility of emotional problems.[1]

It is not even clear how a noise background affects human performance in other tasks. Some experiments have indicated that noise may be advantageous in routine tasks, at least if it is not annoyingly loud.[2] In tasks requiring mental concentration, noise often lowers work quality.

A common experiment in secondary school science classes involves giving students two sets of arithmetic problems. One is worked in a quiet environment, the other while noise is generated from an audio tape player. Performance usually falls dramatically when the noise is turned on. This hardly constitutes a serious scientific experiment, however, because the subjects know the results expected, the noise is not present long enough to allow the students to become used to it, and the Hawthorne effect influences the results.[3]

Noise apparently has an effect on personality. Studies have indicated that men who work in noise surroundings tend to be more aggressive and distrustful. They also tend to have more family problems than normal.

---

[1] After the author gave a talk during which this was mentioned, two of the members of his audience rushed up to the podium to report with great concern that they never dreamt. Were they then prime prospects for emotional problems? The answer, of course, is that most people don't realize they have been dreaming during sleep.

[2] A few years ago, many dentists piped noise through earphones to patients to decrease the pain associated with drilling.

[3] At the Hawthorne iron works, productivity increased when illumination was improved. Some time later, productivity increased again when illumination was decreased. Workers (or students) respond favorably when they feel changes are being made for their benefit.

A 1969 English study of the neighbors of Heathrow Airport (London) showed a higher-than-normal incidence of mental illness.

This rather random collection of thoughts on the psychological effects of noise perhaps emphasizes above all how little is known in this area. Undoubtedly there is a relation between psychological effects and certain of the physiological effects mentioned in the preceding section. Logical decisions about the tolerable noise levels for human beings have to await much more scientific research. Until then, however, there are enough warning signs to indicate that the governmental and public policy should be to reduce and control noise pollution by all actions which do not cause serious economic and social consequences. In the remainder of this chapter, we consider several of the most common noise sources and the corresponding options available to control this part of the environment.

7. Transportation Noise

Transportation is a major source of noise in an urban region. In certain of the New York City subway stations, noise levels exceed 110 dB(A) as the train screeches to a halt. This is a level at which industrial workers would be required to use ear protection with exposure over any significant length of time (Fig. 7-4), but there are no regulations to protect the customer; one can only hope his exposure is sufficiently brief.

The Montreal Metro system demonstrates that subway noise can be reduced by 80% by proper lubrication, reasonably new and well maintained equipment, and rubber wheels instead of steel-on-steel. Noise control can only be obtained at a cost. Rubber wheels are not self-guiding, and additional small wheels rotate laterally to guide the car; steel wheels normally riding above the tracks are used to travel through switches. The cost of this complex system is further increased by the larger energy demands because of more friction. In New York, the subway system is already a major contributor to the energy crisis.

The newer, mass rapid transit systems are notably quieter than the New York system. Ballast in tracks to decrease vibration reduces noise by as much as 10 dB(A), particularly because the tunnels and stations tend to trap the noise pressure waves and cause them to reverberate back and forth. Tracks with uneven joints are another source. Thus, it is not surprising that the new BART system in San Francisco has a quietness which astonishes New York visitors.[1]

Another major source of transportation noise is vehicular traffic, cars and trucks. At a distance of 50 feet from the highway, where vehicles are travelling at 35 miles per hour, noise levels may be:

Automobile	60 dB(A)
Motorcycle	70 dB(A)
Truck	80 dB(A)

---

[1] The most impressive feature of the BART system is the system for automated ticket selling and fare collection. There are no coupons, no tokens, and no waiting in line. A ticket console sells a ticket like a credit card for any amount up to $52.00. The amount is coded magnetically in the card.

As the passenger enters the platform, he slides his card into a console which records the starting point of the trip. When he leaves the platform at his destination, the amount of the fare is automatically subtracted from his card. If there is a balance, the card is returned to him. If the trip uses the entire cash value, the card is "swallowed." If the card value is less than the cost, the turnstile locks, the card is returned, and he must go to an Addfare console where he feeds in the needed amount.

The system is an early version of the "cashless society," in which each person carries a card representing his bank balance and charges are made against this. The intriguing question is how long it will be before some "entrepreneur" builds equipment which will raise the "credit" left on a card for a fee.

10 dB can be added if the vehicle is accelerating, and noise also increases approximately 9 dB for every doubling of speed.

These data are only very approximate. Actual measured noise levels may vary as much as 15 dB depending on the road condition, the maintenance of the vehicle, and the way it is operated. Even the tires have a strong effect on the noise level (Fig. 7-5). Finally, the horn, if it is to serve the purpose of alerting others, has to give an annoying noise.

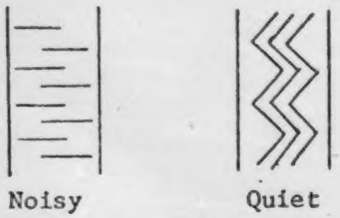


Fig. 7-5 Two common tire treads

Of all the states, only California has made a serious attempt to enforce vehicular noise standards. New York also has a legal restriction, and several other states are planning action.

The last major component of transportation noise is associated with aircraft. The widespread publicity given to the SST (super-sonic transport) debate over the past few years revolved around the two possible environmental impacts: the effects on the ozone in the atmosphere, and the noise problems. The airport noise problem, occasionally so severe that normal conversation is impossible in a home near the end of a runway in use, involves several different groups:

(1) Those who lived there before the airport was built

(2) Those who moved there after the airport, but in a period of smaller and quieter planes

(3) Those who moved there after noise levels reached the present status.

The remarkable feature (at least to someone who lives far from airports) is the large number of people in category (3). Around the new London international airport as well as Dulles in Virginia, housing developments and small cities seem to arise shortly after the airport is in operation.

Airport noise is so annoying to neighbors that major efforts are being made both to monitor noise levels and to design quieter planes. There are now five systems regularly measuring noise. At Kennedy International, LaGuardia, and Newark Airports serving New York City, runway noise is continuously monitored to prevent violations of the federal regulations on allowable levels. (Level can be controlled by adjustment of engine power and the maneuvers immediately after takeoff to turn away from residential areas). The other two monitoring systems, at Orange County and Los Angeles International Airports in California, make measurements both at the runways and in the neighboring community.[1]

8. Other Sources of Urban Noise

In addition to transportation, other sources of urban noise can be divided into three general categories:

Construction

Construction noise, including street excavations, is a principal concern of the noise codes adopted in 1972 in New York and under consideration in other cities. Two regulations are obvious. First, construction can be restricted to hours when most of the inhabitants are awake -- e.g., 7 a.m. to 7 p.m. Second, when jack hammers, pneumatic drills, and other noisy equipment are used, acoustic panels can be required around the equipment to decrease the noise level at nearby sidewalks and inside neighboring buildings.

---

[1] An excellent discussion of noise monitoring problems is given in the article by Edward A. Starr, Measuring Noise Pollution, IEEE Spectrum, June 1972, pp. 18-25.

City services

Rubbish collection has historically been a source of annoying noise, with the classic picture of the sanitation worker slamming the metal can back on the sidewalk. Plastic bags for rubbish containers, quiet plastic bottoms on cans when they must be used, and stringent noise specifications on the collection trucks are leading to significant improvements.

A second city service includes the emergency vehicles with their varied sirens to warn pedestrians and motorists ahead. As indicated in Chapter 4, research is now underway to design more directional sirens in order to minimize the annoyance to people on either side or behind the emergency vehicle.

Homes and offices

While music and entertainment may be a serious factor in hearing loss, the most pervasive source of noise pollution is the air conditioning equipment. Other appliances and machines also contribute. A Stanford Research Institute study showed one apartment where kitchen noise exceeded 100 dB(A) when the fan, dishwasher, and garbage disposal unit were on. Fortunately, with increasing public awareness of noise pollution, there has followed a marked trend toward an emphasis on noise reduction by the manufacturers. The newer urban building codes include specifications on sound insulation.

9. Conclusion

Noise pollution is typical of environmental problems in many ways.

(1) Scientific knowledge is exceedingly meager, particularly when we attempt to learn the effects of the pollution on people.

(2) The public is only beginning to be aware of the possible threat to physical and mental health. Public education in environmental matters is very new, and even the mass communication media have only a few years' experience in reporting environmental technology.[1]

(3) The public has to be educated on the importance and significance of noise. When unusually quiet vacuum cleaners and power lawnmowers have been placed on the market, the reaction of the public has been that these machines must be low-power and ineffective.

(4) Federal and state governments have not placed a high priority on medical or rehabilitation care for those with hearing loss. While the federal government spends about $300 annually for research per prospective heart or cancer patient, only about $1 is spent each year for each potential deafness case. Even the sale of hearing aids is largely unregulated, with anyone able to sell and no advice given to the purchaser. It was only a year ago that a court decision finally required the Veterans' Administration to make public their extensive rating of commercially available hearing aids.

---

[1] There are relatively few books on noise pollution, and acoustic and noise engineering have not been popular careers for young people. Very few colleges give courses in noise measurement or control. One book written for the general public is by Theodore Berland, "The Fight for Quiet," Prentice-Hall, 1970. Since the publication of this book, Berland has written similar,shorter articles for Smithsonian, Environmental Education, and other magazines.

(5) Government has not given high priority to enforcement of anti-noise regulations for public protection. Finally at the end of the 1972 legislative session, the Environmental Protection Agency (EPA) was given the authority to set standards for noise production by transportation equipment and other machinery. Product labels will have to specify noise characteristics. Perhaps most important, in federal purchasing preference will be given to low-noise models.

In spite of this progress during the last few years, urban noise levels continue to rise. The future control of noise pollution depends critically on public understanding of the total problem, including the possible effects on human health, and recognition of the economic and convenience compromises necessary to permit effective action.

CHAPTER 8

EPILOGUE

In the socio-technological problems considered in these notes, and over the much larger group represented by these examples, decisions and actions are forced on the nation. The question is not: should we allow developments which might create certain problems? The difficulties already exist. What do we do next?

Furthermore, a decision must be made in most cases. No decision at all, or postponement, is a decision to do nothing -- to maintain the existing situation. While the example of the approaching hurricane in Chapter 1 (and the question of whether to seed or not) is perhaps unusual in the short time span allowed for a decision, in many ways the seeding question epitomizes the entire range of socio-technological problems. Should we intervene? That is, should we try to change and control the future course of events? If so, how?

Within the scope of the problems of the energy crisis, environmental degradation, drug epidemic, transportation, education, and health services, the technological aspects are only one element. While the human mis-use of technology has created certain of the problems, that same technology is essential if we are to improve the situation. The question is how to use that technology in a way which has the best chance of ameliorating the problem -- of improving the quality of life for the population. The proper decisions and actions depend on coupling a basic understanding of the capabilities of technology with a concern for human and social values.

Two recent developments demonstrate the difficulties we face in making the appropriate decisions in these problem areas.

Artificial kidney machines

As mentioned in Chapter 3, the United States now possesses the technological capability to save the lives of many people who die from acute kidney disease. The artificial kidney machine can be used to cleanse the blood regularly (typically weekly) until the patient's kidneys recover or until a transplant can be arranged. There are two problems: (1) the number of machines is severely limited, so that any hospital with such equipment has to have a committee which restricts access to patients who are otherwise healthy and who have an excellent chance of full recovery; and (2) the costs of treatment average about $20,000 per year.

In the closing days of the 1972 legislative session, Congress passed a bill (H.R.1 covering social security benefits), which included a section providing that Medicare would be extended to encompass treatment of serious kidney diseases for all Americans, regardless of age. During the brief consideration of this provision of the bill, cost estimates were presented varying from $35 million the first year to as high as $250 million after several years. The estimates of the number of lives which might be saved range from 5,000 to 30,000 per year.

By the time the next Congress convened in 1973, more careful studies of the probable cost had been made. In particular, future cost predictions now included the recognition that certain patients would have to be on the machines permanently; hence, the costs would rise steadily to $1 billion per year by 1983.

The response of legislators was predictable. One of the co-sponsors, Senator Burdick of North Dakota, said that he would never have agreed if he had known the cost. The entire program was threatened by this reaction.

Certainly part of the reconsideration could be traced to the question of the desirability of keeping alive indefinitely at public expense the patients who were suffering from a variety of chronic ailments as well as kidney failure. Second, the precedent of government care for the chronically ill, regardless of cost, would certainly be extended to other categories such as hemophilia and muscular dystrophy. The total costs would rise steadily.

The optimum decision in such cases is certainly not clear. What value do we associate with keeping people alive? Where do we set the limit on public sacrifice to maintain life for the chronically ill? As technology evolves to allow more and more life-saving medical care, how are the costs covered?

Illegal immigration

In Chapter 2, we discuss briefly the problem of illegal immigration, which apparently is roughly comparable in size to those entering the country legally. The illegal immigrants pose the additional problems that they are often unskilled or unable to find regular work and that their presence tends to lower the wage scales, at least in some parts of the country. A significant part of the U.S. population growth would disappear in the absence of illegal immigration. The rapid rise in the deportation of illegal immigrants in the last few years indicates an intensified program by the federal government.

In early January, 1973, objections were voiced to the way that immigration officials were carrying out this effort. Supported by four Congressmen, the Coalition of Latin Americans and Friends of Latin

America and the Committee for the Protection of the Foreign Born objected. They claimed that groups of Latin Americans were being stopped in subway stations and questioned. As a result, legal aliens lived in an "atmosphere of fear." The Immigration and Naturalization Service agreed that groups had been questioned, but argued that the areas chosen were known centers for illegal immigrants and that a significant number of illegal aliens had been caught in the dragnet.

The question of exactly where and how to draw the line between protection of individual liberty and protection of public interests is a problem which dominates many of the socio-technological problem areas. Even in the energy crisis, to what extent can restraints be placed on the actions of the individual?

Future effects of today's decisions

Even if we could answer the questions posed by the two, preceding examples, there is another concern: What will be the future effects of the socio-technological decisions which are made today? Can we anticipate, at least to some extent, the future consequences of actions which may ameliorate current problems?

Unfortunately, history is not very comforting. A long list can be given of technological changes which ultimately led to side effects totally unanticipated. The classic example is probably the automobile, which was heralded at the beginning of this century as a savior of the urban environment (then plagued by horse urine and manure).

Other examples are the typewriter, which liberated women from the home by providing a vast number of attractive jobs, or the window screen, which in Latin America was a major factor in the control of malaria. A more recent unfinished example is the data-processing digital computer which has saved the banking industry (the Bank of America has estimated that, without computers, it would now have to employ every adult in the state of California for routine record-keeping). In this case, the long-range effects can only be guessed.

The specialty of trying to anticipate the primary and secondary effects of new technology or of socio-technological decisions has come to be called technology assessment -- a term first popularized in the late 1960's by Congressman Emil Daddario of Connecticut. In 1972, Congress established an Office of Technology Assessment, specifically designed to give advice on proposed new legislation.

Technology assessment really has two facets. The first is the prediction of the direct effects of a new program -- usually the relatively short-term effects. When a weather modification project to increase snowfall in Colorado is started, it is possible to anticipate the problems for the mountain residents and the social and economic changes which result in the southwest as a consequence of the greater availability of fresh water.

It is much more difficult to predict the longer-term and secondary effects. Scientifically we know so little about meteorology, we can not foresee with confidence the ultimate effects (if any) on the weather throughout the rest of the nation or in other countries. From the social sciences, what changes will occur when more water is available to Nevada, Arizona, and California? If parts of the desert are made arable, will new cities rise in desired locations, or will the changes in population patterns cause aggravation of societal problems?

Thus, technology assessment must encompass both scientific prediction and anticipation of social and behavioral changes. Even in the easier scientific futurology (predicting the science-technology developments of the future), the time lag between discovery-invention and widespread use is decreasing so rapidly, we find it difficult to predict beyond a decade into the future.[1] Yet we have seen throughout these notes that most technological change has an inherent time lag of close to ten years (a new power plant takes seven years to build; new auto safety devices are not generally used for seven years just because of the old cars on the road).

The rapidity of change couples with the increasing complexity of every social system (that is, every program or decision has impact on so many social organizations and human activities). In consequence, successful technology assessment is one of our most difficult intellectual challenges.

---

[1] For scientific prediction, a popular approach is the Delphi method. A panel of experts is selected. Each is asked (by mail) to predict the date of a certain scientific or technological innovation. In phase 2, each is provided with the distribution of dates chosen by everyone and allowed to modify his estimate if he wishes. In the next iteration, each is allowed to state briefly his reasons when he disagrees with the concensus. These arguments are distributed to the panel members. The ultimate goal is the best possible set of group predictions.

PROBLEMS

Problem 1-1

The recent President's Conference on Aging concluded that one of the major problems in this country for the elderly (and the physically disabled) is the difficulty of obtaining transportation around their city or town. As a result, many of these people are tied to their homes or apartments. What are some of the problems -- that is, what are some of the situations in a city or large town which unnecessarily cause difficulty in mobility for this population? For example, one answer is that buses usually have to be entered by climbing up from the curb on several narrow and relatively steep steps, whereas it would be quite easy to build buses with the floor at sidewalk level.

Problem 1-2

We desire to design a mobile, automated, multiphasic health testing center for elementary school children. The plan is to park the trailer in front of a school for enough days to allow all the children to be checked, with parental consent and attendance by a parent if he wishes. The trailer will return to the school each year to recheck the children in grades K-4.

The mobile unit is being designed to test children in a heavily populated area of the city where there are a mixture of ethnic groups. We guess that many of the children will not have had a physical examination, or at least none for which records are available.

(a) Describe the tests you would recommend for such a center, and suggest ways in which the tests might be administered or special features to make the program attractive to the customers.

We now turn to some overall planning. We decide to try to cover 14 schools during the year. We anticipate one day per week on the average will be lost because we can not interfere with special events such as aptitude tests, parents' day, and so on. On the average, each school will have 500 children to be examined.

(b) How long should we plan for the full screening of each child?

(c) How should the children be scheduled in the light of our limited space and the conflicting desire to minimize disruption of the regular school work?

(d) How many personnel will we need to employ? Here the maintenance of the equipment can be neglected. You should consider the type of testing, however, and include the fact that most employees can not work for five straight hours without a break.

(e) If typical salaries are paid to the operating personnel, the total annual cost of the center might be three times the total salary figure. Whether to multiply by 3 or 4 or some other number depends on the elaborateness of the equipment, the number of years before the equipment has to be replaced, the cost of maintenance, the cost of ensuring medical follow-up for children who require it, etc. We use 3 as a typical multiplier. What is then the cost per child examined?

(f) Comment on the feasibility and desirability of introducing such a program.

Problem 2-1

The data below indicate the predicted population growth in various sections of the world during the remainder of this century. The information is presented in two ways, by the table and also with the figure looking like a mushroom cloud.

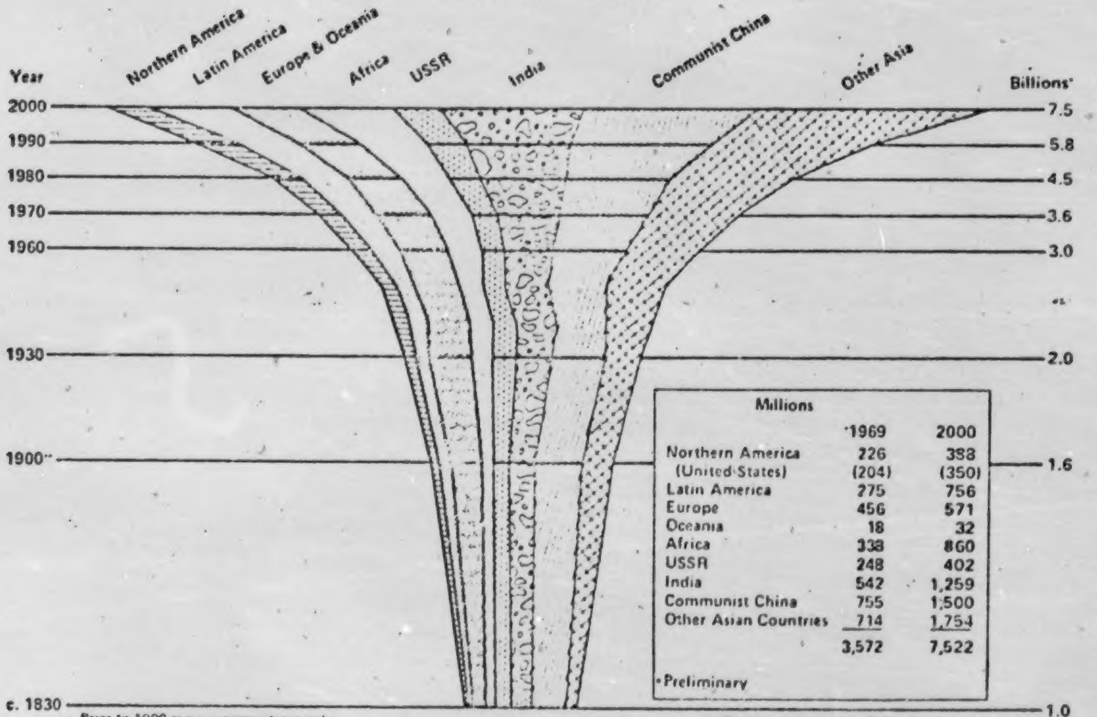
(a) For the section with its population growing most rapidly, what is the annual growth rate, approximately?

(b) Explain briefly the most important reasons for the rather marked differences among the growth rates for the different regions.

(c) If Latin America is to achieve a steady improvement in the standard of living corresponding to a per capita increase of 5% per year in the gross national product (close to the figure representing the current goal of the U.S., so hardly an ambitious objective for this less developed section), at what rate must the GNP increase per year?

The time for the population to double can be measured quite easily from the graph. If we ask when the 1960 population of Africa will double, we place the edge of a small piece of paper on 1960 and mark the width of Africa's population. Sliding the paper over, we mark it again directly adjacent to the first marks. The total width now represents a doubling. Finally, we move the paper upward until this total width represents Africa's population and read the year.

THE POPULATION BOMB
IN THE YEAR 2,000
WHERE THE PEOPLE WILL BE



Millions

	*1969	2000
Northern America	226	388
(United States)	(204)	(350)
Latin America	275	756
Europe	456	571
Oceania	18	32
Africa	338	860
USSR	248	402
India	542	1,259
Communist China	755	1,500
Other Asian Countries	714	1,754
	3,572	7,522

*Preliminary

Prior to 1900 gross appropriations only

Based on assumption of constant fertility levels
Source United Nations Basic Data

Problem 2-2

Huntington is a suburban town just into Suffolk County on Long Island. The population is 180,000, and it has doubled in the past 8 years.

(a) What has been the percent increase per year? Estimate the allocation of this increase among the various causes.

(b) If this rate of increase continues, approximately when will the population reach one million?

(c) What information would you need to predict when the rate of increase of population is likely to start to decrease -- that is, when the population will level out?

Problem 2-3

When we want to study a particular social problem, we often start by trying to find data describing how serious the problem is, in which social groups is it most serious, and so on. In many cases, the data are not in a directly usuable form.

For instance, we might like to know life expectancy in each of the 50 states of the U.S. in order to see where serious health problems seem to be greatest. The only data we can find give the deaths last year per thousand people in the population. Maine is the worst with 11.4, while Alaska was best with 4.9.

(a) Can we dream of any logical explanations for these data? For example, Florida has a high rate, with so many older, retired people living there.

(b) Are these data of any help at all in deciding where there are major health problems?

Problem 2-4

The article by Jean Mayer (footnote on page 79) is a summary of recent national discussion and planning with regard to the problem of hunger in this country -- a topic which has been discussed frequently in very strong political terms.

If we accept the hypothesis that there does exist a national problem, describe

(a) The principal problem or problems.

(b) The primary factors explaining why the problem has not been ameliorated during the past decade with our national prosperity.

Problem 3-1

In an automated health screening center, we desire to give each customer a medical-history questionnaire consisting of 280 queries. There are several alternatives for delivering the questions and collecting the answers:

(1) We might hire personnel from the community and train them to ask the questions in face-to-face interviews.

(2) We can print each question on a separate card and ask the customer to put the card in one of three boxes: Yes, No, Don't Know. Each card stack is then processed automatically.

(3) We can put the questions on audio cassettes in various languages, simply have the customer press one of three buttons for the answer.

(4) We can present the questions in sequence on a video terminal of a computer, again with three response buttons.

(a) What are the major advantages and disadvantages of each alternative?

(b) If you were designing a center for a disadvantaged area of New York City, where most of the customers will have been essentially separated from the health care system, what would be your choice? Explain briefly.

Problem 3-2

The infant mortality among the black population is about 40/1000, among the total population 20/1000. To what extent is the poor U.S. performance the result of the inadequate health care for blacks? In other words, what is the infant-mortality figure for U.S. whites? (Since blacks constitute about 1/9 of the total population, you can assume 9000 live births and find the expected deaths during the first year for the 8,000 whites. Here we neglect the other non-white groups).

Problem 3-3

The importance of increasing hospital rates is related to the duration of the patient's stay for a particular illness or treatment. One study* of five hospitals indicated a remarkable correlation between the length of time patients stay in a hospital and the length of time members of the hospital (nurses in particular) stay in their positions. In the study, five types of patients were considered -- those in for appendectomies, hernias, etc. --, and four classes of nurses. Discuss the factors determining the duration of a patients' stay for a particular ailment. Which of these would you expect to be correlated with the duration of the nurses' employment? A recent census of patients in a New York City municipal hospital revealed a significant number who had been there very much longer than normal. What factors might explain this?

One distinguished expert on health services analysed this study and concluded that the correlation between the patients' stays and the lengths of nurse employment is the result of the efficiency, attitudes and general competency of the hospital administration. "Consequently," a hospital administration which creates a pleasant working environment also leads to longer patient stays -- the patient is not as anxious to get home.

This sort of unjustified argument to explain statistical results in research on social systems is one of the greatest deterrents to meaningful research. Give another example where correlation between two phenomena is simply a result of the fact that both follow as consequences of a third situation or phenomenon, rather than as an indication of a cause-effect relationship between the two. (In the hospital study, a little thought suggests that overloaded hospitals, with their waiting lines of patients to enter, probably discharge patients more rapidly; the turnover rate then overloads the nursing staff, and employees tend to leave for other jobs).

*R.W. Evans, "Standards for Morale: Cause and Effect in Hospitals," Oxford University Press, London, 1964.

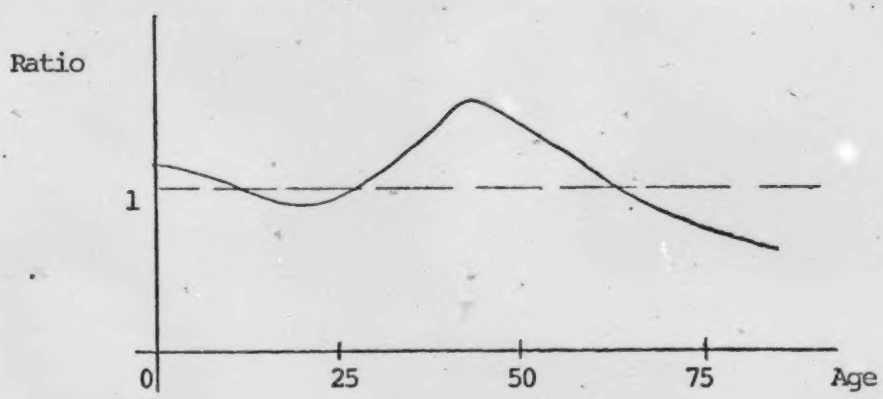
Problem 3-4

The U.S. fetish for dieting and the emphasis on being thin is a relatively recent phenomenon in western civilization. It was not long ago that plumpness was considered a sign of male success and an attractive female characteristic. Middle-aged adults today can still recall the outspoken pride of parents in how "fat and healthy" their children were. What social and medical factors have led to this change of attitude?

Problem 3-5

We saw before that, in the study of any important social problem, we have to be careful to draw intelligent conclusions. If two groups of people both have two characteristics, it does not follow that one characteristic is the result of the other. They both may be results of a different cause, or they may be totally unrelated. This question illustrates the problem in the nutrition area.

We study England and the United States, and we find that U.S. males in middle age have a much higher death rate than their English counterparts. It is traditional that the British like their meat well done, whereas Americans tend to choose medium rare. In Venezuela, where the death rate is even higher than the U.S. for males around 40, a favorite lunch is an egg on raw hamburg (steak tartar).



Mortality rates (males) as a function of age (U.S. compared to United Kingdom).

From this "profound" study of the eating habits of the men of three nations, we deduce that inadequately cooked meat (especially beef) causes diseases which up the death rate alarmingly. Armed with this scientific knowledge, we form an organization devoted to popularizing over-cooked meat. To be successful with such an organization, we have to be able to solicit contributions from wealthy people who are converts to our cause. Unfortunately, these people won't contribute unless they can deduct these contributions when they pay their U.S. income tax (i.e., they don't have to pay tax on the money they give us).

We find that legally such deductions on tax are possible only if our organization is granted what is called a C-3 status by the federal government. We apply for this status to the U.S. Internal Revenue Service (IRS). We claim we are a legitimate, charitable organization working in the interests of the welfare of the American people. Unfortunately, the IRS refuses to grant us C-3 status; indeed, they simply state we are an organization of crackpots, with no legitimate reason for existence.

If you believe our cause is absurd, explain why. What other factors might more logically cause the high U.S. and Venezuelan mortalities of 40 year-old males?

Problem 4-1

One of the attractive alternatives to try to reduce auto accidents is to make driving regulations uniform throughout the United States, rather than to tolerate different laws, even in neighboring states. Because the Constitution leaves with the states all powers not specifically assigned to the federal government, and because there has been a strong public attitude against unnecessary centralization of governmental authority, it has proved difficult to obtain consistent laws among the states. Indeed, the principal power of the federal government is in the disbursing of funds; in the Interstate highway program, billboards were forbidden by the threat to withhold money.

At the present time, there is a major argument over whether to allow a right turn during a red light. Twelve states already permit this, 13 others allow such turns whenever a sign is posted. The National Highway Traffic Safety Administration has proposed a uniform rule always to allow right turns on red unless a sign specifically forbids it. (In every case, the driver turning right must yield the right of way to pedestrians or to traffic moving through the intersection on green).

The most vehement opposition has come from officials in the major eastern cities. What are the motivations and arguments against the NHTSA proposal? What are the principal arguments for? How would you personally vote?

Problem 4-2

Summarize briefly the arguments for and against the air bag as an auto safety device. If you saw the movie made by Allstate, indicate which arguments you would have included against the air bag requirement in order to give an unbiased presentation. If you did not see the movie, list the various arguments for and against, assign relative weights to each of these, and determine a net quantitative argument for or against.

Problem 4-3

Laws in each state differ as to the alcoholic content allowed in the breath or blood before a person is presumed to be guilty. In many cases, a half dozen drinks over two hours (or an equivalent number of bottles of beer) must be consumed before the average adult male reaches the "guilty" level. Since there are perhaps ten million alcoholics and serious, chronic drinkers in the United States, the role of alcohol in auto accidents is apparent. For your state, determine the legal definition of being "under the influence of alcohol," the number of drinks required to reach this state (this depends on the weight of the individual), and the tests which police are allowed to make on a driver they suspect.

Problem 5-1

One of the most terrifying features of the drug problem is the growing number of babies who are addicts at birth because the mother used heroin during pregnancy. In 1972, one of every 40 babies born in New York City was addicted and had to go through withdrawal, with the accompanying convulsions, diarrhea, and vomiting. New York is certainly atypical, since half the U.S. heroin addicts are there, but perhaps that city is a forewarning of problems elsewhere.

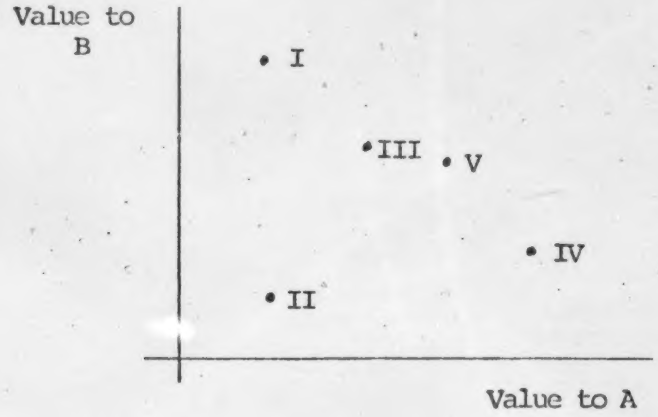
It is only very recently that such infants have been kept alive, so there is no indiction yet as to what the effects will be in later life. Even if the prospective mother could be found for treatment during pregnancy, the problem would not disappear: withdrawal may cause convulsions and death for the fetus, and methadone withdrawal is perhaps worse than heroin for the infant.

If we accept the fact that heroin addiction is not likely to decrease significantly among the young women of child-bearing age, what governmental actions would you recommend to try to reduce the number of newborn addicts?

Problem 5-2

In Pareto decision analysis, we consider each of the groups affected by the decision and the net value of each possible alternative to each of these groups. We then throw out every alternative which is clearly inferior to another. In this way, the complexity of the problem is reduced, and we need only decide among alternatives on the basis of what weight to give each group.

For example, if there are only two groups (A and B), we can plot the values of different decisions.



We show five alternatives (I, II, III, IV, V). We can reject II immediately compared to I since the two have the same value to A, but I is much better for group B. Similarly, V almost dominates III with respect to value to A. Thus, we really have to choose among only I, V, and IV. Which we select now depends on the relative importance of groups A and B.

The parking problem at the university, shopping mall, or downtown center is a decision problem in which Pareto analysis can be applied. Select a specific, familiar example and consider the questions:

What are the interest groups? (The state taxpayers should obviously be included)?

What are all possible options, including combinations of alternatives?

What would be the form of a Pareto model?

Problem 5-3

In a health examination, one of the important measurements is called spirometry. This involves measurement of the volume of air which the man exhales in one breath. This volume is a measurement of the lung capacity and is particularly important in assessing the health of an older person (where lung deterioration and emphysema are apt to be potential problems).

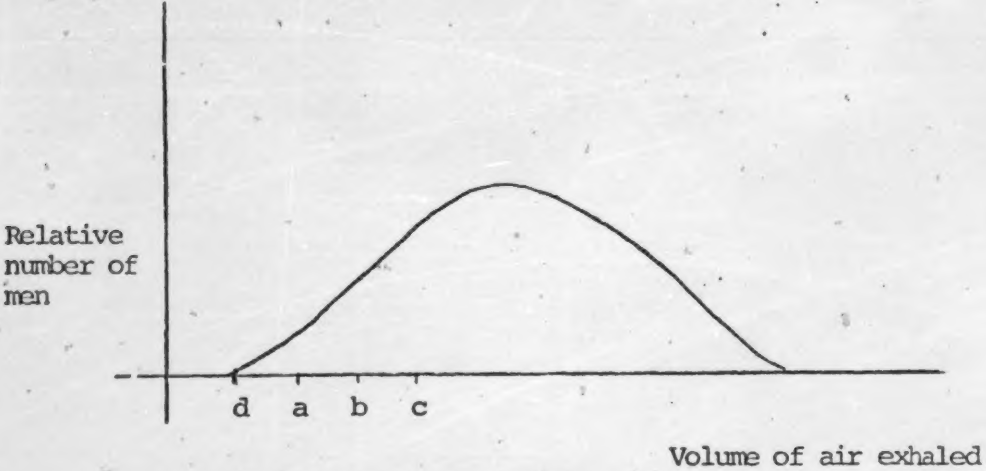
We can make the measurement very simply. The man to be tested takes as deep and full a breath as he can, then exhales as much air as possible into a balloon. The largest distance around the balloon (the circumference) is found with a tape measure and recorded. The measurement should be repeated three times with each person; if the three readings are not within 20% of one another, additional measurements can be made to obtain a measurement in which we can have reasonable confidence.

The spirometry measurement illustrates three important characteristics of most health tests in people:

(1) Success of the test depends on the full cooperation of a relaxed and comfortable subject. If the man being tested is frightened, nervous, or overly tense, test results are of little value. Furthermore, the man must cooperate; no one can make him take as large a breath as possible.

(2) Results for one patient are not precisely repeatable. No one can breathe twice exactly the same way. Consequently we must use an average. In health testing, we often would really like to have this average taken over several days, since in any one day the man's response may be influenced by tiredness, by what he ate the day before, and so on. Normally, however, we have to be satisfied with data collected during one visit to the testing center.

(3) Test results vary markedly from person to person. In many tests we find a distribution of results such as shown in the figure. This variation from person-to-person is apt to be quite large; this is true even among totally healthy people.



Spirometry test results.

The above characteristics are particularly troublesome when we try to interpret the results -- e.g., to determine who needs medical care. If we look at the results of the figure and examine the patients more closely, we might find, for example, that

(a) Around region a, half the people are completely normal and healthy.

(b) Around region b, 60% of the people are normal.

(c) Around c, 70% are normal.

In other words, all the test really shows is the people who should be tested further. As we move toward poind d in the curve, there is increasing probability that the man requires treatment -- although even a subject at a is only a 50/50 proposition.

Thus, test results on people are rarely conclusive -- rather they must be interpreted in terms of probability. If we were running spirometry tests on a large group of people, we would have to decide at what level we should recommend further tests. Should we suggest testing of everyone below b? Clearly this would mean expensive and annoying tests of a large number of people who are healthy. On the other hand, if we set the referral level at a, we have fewer "false alarms", but we also have more "misses" (ill patients who are not identified by the test).

As an experiment, test* all volunteering members of the class or a group of friends and plot the results in the form of a curve. Determine who the heavy smokers are in the group and determine whether these people possess a different average than the non-smoking group. If middle-aged individuals can be tested, the effects of smoking are clearer. (If the entire test group consists of young, healthy adults, the test primarily shows the wide range of results achieved from healthy people).

In a different health test, the following results are achieved from tests on 20,000 people:

Test result	%	Fraction needing care
less than 20	60	2
20 - 22	12	20
22 - 24	10	40
24 - 26	9	60
26 - 28	6	80
28 - 30	3	90

*For this experiment, the equipment needed is a small flacid balloon (which can be partially inflated by the expiration of one breath from a person, but not inflated to the point of bursting) and a tape measure. It is desirable to have extra balloons available in case one bursts.

(1) What should be the referral level for further tests if we are willing to have 50% false alarms?

(2) Under the plan of (1), how many ill people will be overlooked or missed?

(3) What should be the referral level to catch 50% of the ill people? What is the corresponding number of false alarms?

(4) If each referral costs $100 for the additional tests and each ill person missed costs $1000 in later care which could have been avoided by early detection, what is the referral level minimizing the total cost?

(5) What factors other than those considered above might influence our choice of referral level? Discuss the problem of choosing the best level.

Problem 6-1

There have been proposals to solve at once both the energy crisis and the solid waste problem by using the waste as an energy source. Several cities have small plants where the incineration yields steam which is sold for the production of electricity or heating. Baltimore is currently building a facility to handle 1,000 tons of solid waste each day. Even with the sale of steam, the cost of waste disposal will be about the same as landfill.

A more promising prospect is to produce oil by a hydrogenation process from organic wastes. In this process, the waste, sodium carbonate, carbon monoxide, and steam are subjected to high pressure (perhaps 3,000 pounds per square inch) at a temperature over 300°C for an hour. About 1.25 barrels of oil are produced for every ton of waste. The total oil consumption is approximately 5,600 million barrels per year in the United States.

(a) If all 900 million tons of dry organic wastes could be collected throughout the U.S. every year from manure, urban refuse, sewage, industrial and agricultural wastes, what percentage of the nation's oil needs could be satisfied?

(b) In practice, only 15% of these wastes can be easily collected. Why?

(c) What percentage of the nation's oil needs might be economically satisfied by hydrogenation of organic wastes?

(A fascinating study of the problems promised by future growth is reported in the paperback book by D.H. Meadows, D.L. Meadows, J. Randers, and W.W. Behrens, III, "The Limits to Growth," Potomac Assoc., Washington, D.C., 1972. While the premises for this computer model study by the Club of Rome have been attacked, the important contribution is the dramatic emphasis on the urgency of such modelling research).

Problem 6-2

The water supply for a large city is an example of resource depletion, even though the supply is continually replenished. When the use rate exceeds the reservior-refillment rate, the resource will disappear.

New York City, with its 8 million inhabitants, uses 1,200 million gallons of water per day. Perhaps 20% of this is wasted in leaks: the two tunnels bringing water to the city are 60 and 50 years old, and the valves have never been closed. Indeed, parts of the underground distribution system are uncharted, at least in available records.

(a) Ever since the drought of the mid-1960's, restaurant owners have contributed to the water-conservation program by not serving a glass of water to a customer unless he asks. Does this policy result in a significant saving? Why was the program started at the time of the drought?

(b) If a reservior toilet is used (with the water tank behind the seat), each flushing uses about 5 gallons of water. From measurements of a toilet reservior, show this fact. Determine the water per person used each day in flushing of toilets (actually the average number of flushes per day is not a known scientific fact, and estimates have varied widely). In the early days of television, the audience for a show was measured by the municipal water company: when a commercial came on, the abrupt surge in water usage was taken as an indication of the number of people watching.

(c) Water conservationists urge homeowners to place a brick in each toilet tank. The flushing pressure, dependent on water depth, is unaffected, but the quantity of water used is reduced. Estimate the percent saving in daily consumption if everyone used a brick. (The program is enthusiastically supported by the brick manufacturers).

(d) New York City is now constructing a third tunnel, both to increase the water supply and to have protection against sabotage or other failure of one of the two existing tunnels. Describe briefly the effects of blockage of one tunnel.

Problem 6-3

In the 1920's, coal dominated energy use in the United States. By 1948, coal provided less than half the energy, and today it amounts for only 20%. Oil has risen slowly from 1/3 in 1946 to 43% now. Natural gas use is increasing at the rapid rate of 7% per year; in 1946, gas gave 14% of the energy, today it provides 33%.

(a) Discuss the reasons for the decline in the role of coal.

(b) What have been the social and economic effects of this decline, particularly in an area such as Appalachia?

(c) If coal gasification and coal cleaning technologies can be developed and are economical, what changes can be anticipated after the 1980's?

Problem 7-1

In recent years students at various colleges and universities have made studies of the noise environment. Some of the worst offenders are subways, aircraft near airports, garbage trucks, construction equipment, city traffic, home lawn mowers, and hi-fi sets. Two approaches which might be employed to reduce noise are regulation and education.

(a) If a noise level meter is available (an inexpensive one can be purchased for less than $50), measure the noise level at various places in your normal environment.

(b) Which sources of noise could be reduced at little or no cost or inconvenience? In each case, can the reduction be achieved by regulation or education?

(c) What people within your acquaintance lead lives which bring their noise environment close to the federal limits for industry?

Problem 7-2

One problem in providing access to urban centers is that a large percentage of suburban commuters travel alone to the city in automobiles. One suggestion is to reserve a lane of expressways for cars with two or more occupants. The hope is to encourage car pools.

(a) What arguments might commuter groups give against this plan?

(b) What problems of enforcement will be faced by the traffic department? (When this was tried in 1972 in San Francisco, some drivers carried mannequins on the front seat -- hopefully with lap belts fastened).

(c) What effects would you anticipate in the noise environment if this plan were successful?

Problem 8-1

One of the serious problems we face comes from the natural tendency to seek simple solutions for problems which are actually exceedingly complex. Often these obvious "solutions" have side effects which are more serious than the problem we originally wanted to solve.

For example, there are strong forces arguing for the replacement of phosphates in detergents by washing soda (largely because of the environmental effects of phosphates). What are some of the possible side effects if this change is successfully made?

For one thing, in 1970 at the various poison control centers in the U.S., there were about 4,000 people treated who had swallowed household detergents. No one knows how many more people took care of themselves or were treated by their personal physicians. If soda had been substituted for the phosphates, these people might have been very seriously injured (for example, with severe narrowing of the throat passage as a result of the lye).

If you were director of the U.S. Environmental Protection Administration, discuss how you would develop a model which would guide your decision on evaluating the human poisoning effects of detergents.

Problem 8-2

A table of letter frequencies in normal English writing is given at the end of this problem. From these data, we can evaluate the operation of the conventional typewriter keyboard which was designed around 1872. The widespread use of the typewriter in this century has been one of the major contributions to the liberation of women through the creation of millions of secretarial jobs. In China, for example, typing is much less common because 5,000 symbols are required (the typist selects one of an array of plates, slides this into the typewriter, and then forms the desired symbol). The original design, the one we now use, was chosen to prevent the keys from jamming as they approached the roller from various directions.

There are many common inefficiencies in the present arrangement. Letters like A and E and D are difficult to hit, unusual letters like J and K are among the easiest to type.

In 1932, August Dvorak developed a vastly improved keyboard, which is probably close to the optimum. Actually, the best design has never been determined since what is best depends on the criteria used (do we want to divide the words evenly between the two hands, minimize the times when two successive letters requiring the same hand or finger appear, minimize the total distance travelled by the two hands in a day's typing, and so on?)

The Dvorak keyboard has the three rows of letters from left to right:

? , . P Y F G C R L /
A O E U I D H T N S -
, Q J K X B M W V Z

This Dvorak Simplified Keyboard (DSK), available on regular typewriters, has several important characteristics:

(1) In an average day, a typist's fingertips move something like 15 miles with the standard keyboard (called the QWERTY system) compared to about one mile with the DSK.

(2) Many typists can double their speed and reduce errors by switching to the DSK.

(3) Fatigue is reduced.

(4) Learning is much simpler and faster.

In addition, other quantitative characteristics can be calculated from the table of letter frequencies:

(a) With each keyboard, determine the probability of each of the eight fingers being used. Here we omit numbers and punctuation marks.

(b) With each keyboard, compare the probabilities that a desired key is in each of the three rows.

(c) With each keyboard, compare the probabilities of using the right hand.

(d) If we introduced DSK typewriters into our offices, the typists would have to learn a new system, which typically requires a few weeks. Comment on the possibility that then the typist could easily switch back and forth between DSK and QWERTY typewriters.

Several articles have appeared recently in the popular press on the inefficiency of the standard typewriter keyboard. One is the "Tyranny of QWERTY," by Charles Lekberg in Saturday Review, Science, October 1972, pages 37-40. For readers unfamiliar with probability concepts, a great introduction is provided by Warren Weaver's paperback, "Lady Luck," Doubleday Science Study Series, 1963.

Probability of letters (out of 1000)

E	132	R	68	L	34	G	20	V	9
T	104	I	63	F	29	Y	20	K	4
A	82	S	61	C	27	P	20	X	1
O	80	H	53	M	25	W	19	J	1
N	71	D	38	U	24	B	14	Q	1
								Z	1

INDEX

Accidents in U.S., 96-139
  automobile, 97
  fire, 134
Active restraints in autos, 104
African bees, 5
Air bags, 106, 108, 112
Alcohol and driving, 117-126, P 4-3
  countermeasures for, 124-126
  legal measures in, 119
  technology for, 121
Amplitude, 211, 214
Appropriation bill, 76, 94
Arrow, K., 182 (footnote)
Assessment, technology, 236
Audiometer, 213, 223 (footnote)
Authorization bill, 76, 91
Auto accidents, 96
  collision one in, 98
  collision two in, 101
  collision three in, 105
Automated health testing, P 1-2, P 3-1
Auto safety controversy, 105
Auto safety regulations, 102 (table)

Babies, unwanted, 49
BART system, 227
Bees, African, 5
Bell, Alexander Graham, 215 (footnote)
Buffalo, harvesting of, 27
  population of, 22-27
Building codes for fire, 134

Central-city health problems, 68, 74
Cloud seeding, 9
Coal usage, P 6-3
Cocktail party effect, 220
Collision one, 98
Collision two, 101
Collision three, 105
Commons, tragedy of, 178
Cook, Earl, 189 (footnote)
Cost of living, 57
Counter-intuitive behavior, 141
Criterion, 156

dB, 215
dB(A) scale, 216
deciBels, 215

Decision, 139-182
  competitive, 170
  multidimensional, 145
  multi-stage, 161
Decision analysis, Pareto, P 5-2
Decision theory, 179 (footnote), P 5-3
Delphi method, 238 (footnote)
DES additive, 84
Doubling rule, 37
Driving while intoxicated (DWI), P 4-3
Drug addiction, infant, P 5-1
Drug control, 140-150
Dynamic model, 5, 27

Election-night predictions, 157
Electric energy, 193
Emergency medical service, 126
  breakdown in, 127
  call for help in, 128
Energy, 184
  electrical, 184, 193
  from fossil fuels, 204
  geothermal, 203
  kinetic, 184
  of man, 185
  from natural resources, 187
  nuclear, 194
  potential, 184
  solar, 205
  from solid waste, P 6-1
Energy crisis, 183-208
Energy research, 202 (table)
Energy use, 188-194
Environmental effect from energy, 187-192
Exponential, plotting of, 38

Family size in U.S., 45
Fire accidents, 133
Flammable Fabrics Program, 134
Forrester, Jay, 141
Fossil fuels, 204
Fourier theorem, 212
Frequency, 211

Game theory, 171-179
Geothermal energy, 203
Growth, exponential, 35, 38-44
  sigmoid, 39

Hardin, G., 178

Harvesting of buffalo, 27
Health delivery management, 85-87
  in hospitals, 85
Health service system, 56
  current performance of, 58
  doctors in, 66
  management of, 85-87
  for non-whites, 68
Health test, automated, P 3-1, P 1-2
Hearing capability, 209
Hertz, Heinrich R., 212 (footnote)
Horsepower, 191
Hospital management, 85
Hurricane seeding, 1-3

Ice cream wagon problem, 175
Immigration, 46-49, 235
Inertia in population, 16
Infant addiction, P 5-1
Infant mortality, 64, P 3-2
  in other countries, 65 (figure)

Joule, James, 184

Keyfitz, Nathan, 29
Kidney machine, 86, 234
Kinetic energy, 184

Lamprey, sea, 19
Lead poisoning, 74
  from paint, 75
  from gasoline, 77
Life expectancy, 59
  male, 63 (table)
  by state, P 2-3
  in U.S., 60 (figure)
  world, 63 (table)
Logarithmic scale, 43

Mayer, Jean, 79 (footnote)
Meyer, James, 2 (footnote)
Model, 5
Montreal Metro system, 226
Morganstern, Oskar, 171
Mortality, infant, P 3-2
Multi-stage decision, 161

von Neumann, John, 171 (footnote)
Noise, 209-232
  construction, 229
  home, 230

Noise,
  measurement of, 210
  physiological effects of, 221-224
  psychological effects of, 224-226
  from rubbish collection, 230
Noise duration, 217
Noise environment, 209-232
Noise levels, 218 (figure)
Non-zero-sum game, 175-179
Nuclear energy, 194
Nutrition, 79-83

Office of Management and Budget, 90
Office of Science and Technology, 90
OMB, 90
Orthotics, 89

Pareto decision analysis, P 5-2
Passive restraints in automobiles, 104, 109
Payoff matrix, 172 (figure)
Population, 16
  of buffalo, 22
  of elderly in U.S., 88
  female, 29-33
  in Indonesia, 16
  town, P 2-2
  in U.S., 28
  of world, P 2-1
Population control, 45-51
Population model, 28-33
  use of, 33
Population predictions, 28-33
Potential energy, 184
Potomac River ecology project, 166-170
Precipitation control, 8
Prosthetics, 89

Reaction time, 123
RECAT report, 107-116
Resource depletion, 187
Resource management model, 27
Route selection, 158-164
Rubbish collection, 230

Schmidt-Nielsen, Knut, 186
Sea Lamprey in the Great Lakes, 19
Seat belts, 103
  for children, 104
Seeding, 1, 9, 13
  of clouds, 9
  of hurricanes, 1
Shamir, Uri, 179 (footnote)

Shapley analysis, 169 (footnote)
Sigmoid growth, 39
Sinusoidal signal, 212
Skywater, Project, 10
Solar energy, 205
Solid waste, 41-43
  and energy, P 6-1
Speed, of light, 211 (footnote)
  of sound, 211 (footnote)
Spirometry, P 5-3
Stormfury, Project, 8
System Studies, 4

Tragedy of commons, 178
Technology assessment, 236
Transportation for elderly, P 1-1
Transportation noise, 226
  of tires, 228
  of aircraft, 228
Tribus, Myron, 151

Unwanted babies, 49

Venereal disease, 78

Walsh-Healey noise limits, 219
Warren, Richard M. and Roslyn P., 209 (footnote)
Waste, solid, P 6-1
Water usage, P 6-2
Watt, James, 191 (footnote)
Weather modification, 1-13
  to control natural disasters, 13
  inadvertent, 12
Widget problem, 151-155
Williams, J. D., 173
World population, P 2-1

Zero-sum game, 171-175